ABSTRACT

Title of dissertation: MODELING PLASMA TRANSPORT AND WAVE GENERATION DURING IONOSPHERIC MODIFICATION EXPERIMENTS

Aram Vartanyan, Doctor of Philosophy, 2015

Dissertation directed by: Professor Konstantinos Papadopoulos
Department of Physics and Astronomy

Research on ionospheric modifications led to numerous studies of practical importance. This thesis highlights work carried out on this subject. Following an overview of plasma physics and ionospheric modification in Chapter 1, three main topics are discussed.

Chapter 2 examines the effects of long-term HF heating of the ionosphere. It was found that the plasma expands and transports along Earth's magnetic field line, resulting in the formation of plasma "tubes" referred to as artificial ionospheric ducts. While a computational model of HF-heated plasma transport was previously presented, experimental observations of artificial ducts and comparison against the model were missing. A study was conducted by performing several ionospheric heating experiments at the HAARP facility during different ionospheric conditions and times of day, and recording their effects with instruments on-board overflying satellites. The work culminated in the first large collection of satellite observations of HF-induced ionospheric ducts and plasma transport. Modeling comparisons against

a representative subset of observations established the basic physics picture of ducts and their physical characteristics. Moreover, we present the first observations of HF wave focusing by ionospheric density depletions.

Chapter 3 deals with the study of Ionospheric Current Drive (ICD), a method for generating Ultra Low Frequency (ULF) and Extreme Low Frequency (ELF) waves using modulated heating of the F region ionosphere. A wave generation/propagation model of ICD had been presented along with successful proof of concept experiments. However test comparisons of observations and theory had not been performed. To this end, we carried out a set of parameter-sweep simulation studies that reveal the conditions for maximum generation efficiency. With these considerations, we show that the frequency dependence of the generated wave amplitudes predicted by the model is in qualitative agreement with the proof of concept experimental results. Future work will be necessary to fully understand the frequency response of ICD generated waves.

Chapter 4 deals with the generation of Very Low Frequency (VLF) electromagnetic ("whistler") waves. Artificial whistler waves were known to be generated by modulated heating of the polar electrojet. We present the very first clear experimental observations of whistler waves generated by continuous heating of the upper ionosphere. The proposed generation mechanism relies on parametrically excited Lower-Hybrid (LH) waves and their non-linear conversion to whistler waves. After generalizing an existing LH-whistler conversion model, we present simulation results that are in good agreement with the observed whistler wave spectrum. A combination of the qualitative discussions and simulation results explains major peculiarities

observed in the spectrum.

With regard to the first topic, the author took part in experimental planning, processed satellite data, performed simulation runs of an adopted plasma transport computational model, and compared simulation results with satellite observations. In the second topic the author adopted the ICD generation computational model, performed parameter-sweep simulation runs, and compared simulation results with experimental observations. In the last topic, the author took part in experimental planning, processed satellite data, wrote and benchmarked a computational code for an adopted mathematical model of LH-whistler conversion, performed simulation runs after generalizing the computational model, and compared simulation results against satellite observations.

MODELING PLASMA TRANSPORT AND WAVE GENERATION DURING IONOSPHERIC MODIFICATION EXPERIMENTS

by

Aram Vartanyan

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor Konstantinos Papadopoulos, Chair/Advisor
Professor Adil Hassam
Professor Phillip Sprangle
Professor Douglas C. Hamilton
Professor Raymond Sedwick

To my greatest inspirations: mother *Gayane*,
father *Armen*, and sister *Nune*.

To my loving Grandma, *Zinaida*,
the kindest person I will ever know.

Acknowledgments

I owe a great deal of gratitude to many people who have made this dissertation possible and have made my graduate experience one that I will never forget.

I would like to thank my advisor, Professor Dennis Papadopoulos, for giving me the unique opportunity to work on challenging and interesting projects during my time with the Space Plasma Physics Group. The technical way of thinking and approach to problem solving that I have gained, in addition to the wide breadth of knowledge and experience, will be something I carry with me and use throughout my life. It has been a pleasure to work with and learn from such a renowned individual.

I would also like to thank my group members and colleagues Bengt Eliasson, Chris Najmi, Surja Sharma, Xi Shao, Joseph Huba, and Michel Parrot for several insightful discussions and fruitful collaborations. A special thanks goes to my mentor Gennady Milikh, for without his extraordinary level of attentiveness and caring this dissertation would have been impossible. Gennady has been like a second advisor to me; and if not for my frequent visits to his office, countless useful discussions, and the patient holding of my hand throughout the beginning years, then my progress would have been much slower indeed.

Much credit is due to the Department of Physics faculty. Participating in the courses taught by our brilliant professors throughout my time at the University of Maryland has been one of my biggest joys. I should also acknowledge help and support from the department staff members, especially Jane Hessing, who is the most caring administrator I have ever met and made bureaucratic processes as painless as

possible. I would also like to thank Professors Adil Hassam, Phillip Sprangle, Douglas C. Hamilton, and Raymond Sedwick for agreeing to serve on my dissertation committee and for sparing their invaluable time reviewing the manuscript.

Many thanks to my friends from home and those throughout my undergraduate and graduate years at the University of Maryland. They have been one of the greatest sources of happiness, several of which have been like family to me. A special thanks goes to one my longest and most supportive friendships with the twins, Anish and Nitin Sydney, to whom I say they *are* family.

Lastly, I owe my deepest gratitude to my family members, who have always been the most important people in my life. Words alone cannot express how much I owe to my mother, father, and sister, who have always stood by me, guided me through my life, and pulled me through the hardest of times. My parents risked everything to bring me and my sister to the United States for an opportunity at a better education and better life, and I couldn't be happier. My sister has been my best friend, biggest source of support, and a model for what it means to be a great person.

It is impossible to remember all, and I apologize to those I have inadvertently left out. Thank you everyone!

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BRIOCHE	Basic Research on IOnospheric CHaracteristics and Effects
CW	Continuous Wave
CPDR	Cold Plasma Dispersion Relation
DEMETER	Detection of ElectroMagnetic Emissions Transmitted from Earthquake Regions
DM	Downshifted Maximum
DMSP	Defense Meteorological Satellite Program
ELF	Extremely Low Frequency
EM	ElectroMagnetic
EMHD	Electron MagnetoHydroDynamics
ERP	Effective Radiated Power
ES	ElectroStatic
FFT	Fast Fourier Transform
FWHM	Full Width at Half Maximum
HAARP	High-frequency Active Auroral Research Program
HF	High Frequency
IAP	Instrument dAnalyse du Plasma
ICE	Instrument Champ Electrique
ICD	Ionospheric Current Drive
ISL	Instrument Sonde de Langmuir
LEO	Low Earth Orbit
LH	Lower Hybrid
LHS	Left Hand Side
MHD	MagnetoHydroDynamics
MS	MagnetoSonic
MZ	Magnetic Zenith
PEJ	Polar ElectroJet
RBR	Radiation Belt Remediation
RHS	Right Hand Side
RK4	Runge-Kutta method of the fourth order
SA	Shear Alfvén
SAMI2	Sami2 is Another Model of the Ionosphere
SEE	Stimulated Electromagnetic Emission
UH	Upper Hybrid
ULF	Ultra Low Frequency
UM	Upshifted Maximum
UT	Universal Time
VLF	Very Low Frequency

Chapter 1

Introduction

Experiences as routine as the melting of ice into water or the boiling of water into vapor convey the notion that matter comes in the familiar solid, liquid, and gaseous states. However, subjecting a gaseous mixture to a large energy source strips electrons from their host atoms, leaving the atoms ionized, and giving rise to a fourth state of matter. Whether the energy source is ionizing radiation, electrical discharge, or bulk gas heating, the result is an ensemble of electrons and ions referred to as a *plasma*. The scientific discipline that developed to study it, *plasma physics*, has a rich history dating back to American chemist and physicist Irving Langmuir, who first described an ionized gas as a plasma in 1927, and discovered that plasmas exhibit periodic variations of the electron density, now known as *Langmuir waves*.

The development of radio broadcasting lead to the discovery of the earth's *ionosphere* – a partially ionized region of the upper atmosphere. The discovery was facilitated by (among other things) the first establishment of transatlantic radio communication, which, considering the earth's curvature, could only be explained by the existence of a conductive layer that reflects radio waves. Scientists soon rec-

ognized that radio wave propagation in the ionosphere was made quite complicated by the plasma being nonuniform and the wave propagation speed being dependent on wave polarization and orientation relative to the earth's magnetic field. To understand these new effects, various scientists, such as E. V. Appleton and K. G. Budden, developed the theory of electromagnetic wave propagation through nonuniform magnetized plasmas. The later development of satellite communication technology and navigation systems required ground-to-space signals to pass through the ionosphere, thus substantially increasing the importance of ionospheric research.

Astrophysicists quickly recognized that most of the universe is in the plasma state, and thus understanding astrophysical phenomena requires a deeper understanding of plasma physics. This field took a major leap due to the work of Hannes Alfvén, who around 1940 developed the theory of *magnetohydrodynamics* (MHD), in which he essentially treated the plasma as a single conducting fluid. This theory has been both widely and successfully employed to investigate sunspots, solar flares, the solar wind, star formation, accretion disks, and a host of other topics in astrophysics.

The launch of the U.S. Explorer satellite lead to James A. Van Allen's discovery of the Van Allen radiation belts surrounding the earth in 1958. The discovery of the radiation belts marked the beginning of the systematic exploration of the earth's magnetosphere and opened up the field of *space plasma physics*.

1.1 Overview

This dissertation concentrates on exploring phenomena associated with ionospheric plasmas. Since the early developments of ionospheric physics, the ionosphere has become an "open plasma laboratory" where scientists can carry out experiments and compare the experimental results against theoretical models. As is discussed later in the chapter, ionospheric "heaters" were developed to probe the ionosphere by injecting High-Frequency (HF) radio waves into the ionospheric plasma, modifying it due to the absorption of the HF energy into the plasma. This allowed a transition from previously *passive* experimental methods to much more effective *active* experiments. These heaters are supplemented by ground and space based diagnostic instruments, such as satellites, radars, receivers and magnetometers, and several others. The injection of HF radio energy into the ionospheric plasma has a number of important effects, three of which will be the salient topics of this dissertation. Following a basic introduction to plasma physics and the ionosphere, three main topics will be discussed in the subsequent chapters. Namely, Chapter 2 examines field-aligned plasma transport due to long-term HF heating of the ionosphere; Chapter 3 deals with generation of Ultra Low Frequency (ULF) and Extremely Low Frequency (ELF) waves by modulated heating of the F region ionosphere; and Chapter 4 deals with the generation of whistler waves by continuous heating of the F region ionosphere.

1.2 Plasma Physics Basics

This section introduces the basic notion of what a plasma is and the mathematical description of plasma dynamics that plasma physicists have at their disposal. While the material presented here is standard and well known by a practitioner, the author feels that its inclusion gives a sense of completeness and a segue to latter parts of the dissertation.

1.2.1 The Notion of a Plasma

The characterizing properties of a plasma are somewhat subtle since ionization of a gas alone is not quite sufficient for it to be called a plasma. The most succinct way to qualify this difference is to introduce the notion of *collective behavior*, which can be understood by considering the forces acting on a neutral particle from an ordinary gas vs. a charged particle from a plasma. A neutral particle moves undisturbed until it comes within an atomic radius of another neutral, resulting in a "hard sphere" collision. So, all of gas dynamics is mediated by these *local collisions*. However, the analogous picture in a plasma is quite different because the motion of plasma particles creates local concentrations of positive or negative charge, giving rise to electric fields. Moreover, the motion of charges constitutes a current, and hence generates magnetic fields that affect the motion of other charged particles *far away*. So, the motion of the plasma not only depends on local conditions, but on the state of the plasma in remote regions as well; and this is what we refer to as collective behavior. Moreover, while it takes a remarkably small fraction of ioniza-

tion for a gas to be called a plasma, a sufficiently large electron-neutral collision frequency prevents plasma dynamics from taking place. If ω_p is the frequency of plasma oscillations and ν is the mean frequency of electron-neutral collisions, then we require $\omega > \nu$ for the ionized gas to behave like a plasma.

Another important property of a plasma - *quasineutrality* - states that on a large scale the plasma is overall neutral, but can have local charge separation on a length scale called the *Debye length*, λ_D; any charge build-up is "shielded" over this length scale (Debye shielding). Thus the mathematical requirement for a plasma system of size L to be quasineutral is $\lambda_D << L$.

Debye shielding is only a statistically valid concept if there are enough particles in the charge cloud, so we require the number of particles in a "Debye sphere", N_D, to be much much greater than unity: $N_D >>> 1$. The above three inequalities must be satisfied for an ionized gas to be properly called a plasma [12, Ch. 1].

1.2.2 Governing Equations

The equations governing plasma dynamics, from a fundamental point of view, are completely known. For simplicity let us assume a fully ionized plasma with electrons and ions of mass m_e, m_i and charges $-e$, e, respectively. The equation of motion for the j^{th} electron and ion is given by the Lorentz force:

$$m_e \frac{d\mathbf{v}_e^{(j)}(t)}{dt} = -e(\mathbf{E}(\mathbf{x}(t), t) + \mathbf{v}_e^{(j)}(t) \times \mathbf{B}(\mathbf{x}(t), t)) \tag{1.1}$$

$$m_i \frac{d\mathbf{v}_i^{(j)}(t)}{dt} = -e(\mathbf{E}(\mathbf{x}(t), t) + \mathbf{v}_i^{(j)}(t) \times \mathbf{B}(\mathbf{x}(t), t)), \tag{1.2}$$

where $\mathbf{E}(\mathbf{x},t)$ and $\mathbf{B}(\mathbf{x},t)$ are the electric and magnetic field vectors. By introducing $\mathbf{E}$ and $\mathbf{B}$, we need to provide equations for their time evolution; these are given by *Maxwell's Equations*:

$$\text{(Faraday's law)} \qquad \boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.3a}$$

$$\text{(Ampère-Maxwell law)} \qquad \boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} \tag{1.3b}$$

$$\text{(Gauss's law)} \qquad \boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \tag{1.3c}$$

$$\text{(no monopoles)} \qquad \boldsymbol{\nabla} \cdot \mathbf{B} = 0, \tag{1.3d}$$

where ϵ_0 (μ_0) are the permittivity (permeability) of free space, $c = 1/\sqrt{\epsilon_0 \mu_0}$ is the speed of light, and the charge/current density, $\rho/\mathbf{J}$, are given by a sum over each particle:

$$\rho(\mathbf{x},t) = e\sum_j \left[\delta^3(\mathbf{x} - \mathbf{x}_i^{(j)}(t)) - \delta^3(\mathbf{x} - \mathbf{x}_e^{(j)}(t))\right] \tag{1.4}$$

$$\mathbf{J}(\mathbf{x},t) = e\sum_j \left[\delta^3(\mathbf{x} - \mathbf{x}_i^{(j)}(t))\mathbf{v}_i^{(j)}(t) - \delta^3(\mathbf{x} - \mathbf{x}_e^{(j)}(t))\mathbf{v}_e^{(j)}(t)\right], \tag{1.5}$$

where $\delta^3(.)$ is the three dimensional delta function, and the positions and velocities (omitting the sub/super scripts) are obviously related by $\mathbf{v}(t) = d\mathbf{x}(t)/dt$. The above set of coupled equations is an exact multi-particle description of a fully ionized plasma. Unfortunately, such an approach to plasma physics is completely intractable. A typical plasma system can contain an Avogadro number of particles, and tracking the trajectory and effect of each particle is impossible. Even *if* it was possible, knowing the trajectory of each particle is a preposterously large amount

of undiscerning information. With that much information about the microscopic dynamics, it would be natural to start taking averages of quantities to have an idea of what the dynamics looks like on the macroscopic scale. Fortunately, *statistical mechanics* tells us that finding such averages is possible without knowing the detailed dynamics of each particle. This statistical description, referred to as the *kinetic theory of plasmas*, is very important and prevalent, particularly in plasma systems where particle collisions are nearly absent - so-called *collisionless plasmas*. The cornerstone of kinetic theory is the *distribution function* - the statistical distribution of particle positions and velocities. Macroscopic quantities, such as number density, flow velocity, and temperature can be extracted by *taking moments* of the distribution function. The distribution function for species α, $f_\alpha = f_\alpha(\mathbf{x}, \mathbf{v}, t)$, with charge q_α and mass m_α satisfies the *Boltzmann equation*

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha + \frac{q_\alpha}{m_\alpha}(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_\alpha}{\partial \mathbf{v}} = C_\alpha, \tag{1.6}$$

where C_α is the *collision operator* that describes how the time rate of change of the distribution function is affected by collisions. The number density. flow velocity, and temperature are respectively given by the 0^{th}, 1^{st}, and 2^{nd} moments of the distribution function:

$$n_\alpha(\mathbf{x}, t) = \int f_\alpha(\mathbf{x}, \mathbf{v}, t)\ d^3v \tag{1.7a}$$

$$\mathbf{u}_\alpha(\mathbf{x}, t) = \frac{1}{n_\alpha} \int \mathbf{v} f_\alpha(\mathbf{x}, \mathbf{v}, t)\ d^3v \equiv \langle \mathbf{v} \rangle_\alpha\,, \tag{1.7b}$$

$$T_\alpha(\mathbf{x}, t) = \frac{1}{3} m_\alpha \left\langle (\mathbf{v} - \mathbf{u}_\alpha)^2 \right\rangle_\alpha . \tag{1.7c}$$

The *fluid model* of plasma dynamics, which is used extensively in the subsequent chapters, can be derived by *taking moments of the Boltzmann equation*. This is a well known textbook exercise that has been reproduced by several authors (e.g. [8, 12, 55, 48]). This procedure yields the *fluid equations*, which are respectively composed of the continuity, momentum, and energy equations:

$$\frac{\partial n_\alpha}{\partial t} + \boldsymbol{\nabla} \cdot (n_\alpha \mathbf{u}_\alpha) = 0, \tag{1.8a}$$

$$m_\alpha n_\alpha \left(\frac{\partial}{\partial t} + \mathbf{u}_\alpha \cdot \boldsymbol{\nabla} \right) \mathbf{u}_\alpha = -\boldsymbol{\nabla} p_\alpha + q_\alpha n_\alpha (\mathbf{E} + \mathbf{u}_\alpha \times \mathbf{B}) - \sum_\beta \mathbf{R}_{\alpha\beta}, \tag{1.8b}$$

$$\frac{3}{2} n_\alpha \left(\frac{\partial}{\partial t} + \mathbf{u}_\alpha \cdot \boldsymbol{\nabla} \right) T_\alpha + n_\alpha T_\alpha \boldsymbol{\nabla} \cdot \mathbf{u}_\alpha = \boldsymbol{\nabla} \cdot (\kappa_\alpha \boldsymbol{\nabla} T_\alpha) - \mathbf{u}_\alpha \cdot \sum_\beta \mathbf{R}_{\alpha\beta} - \sum_\beta Q_{\alpha\beta}. \tag{1.8c}$$

For completeness and the convenience of the reader, a discussion of kinetic theory and the derivation of the above fluid equation from the Boltzmann equation are reproduced in Appendix A, and the various quantities appearing in (1.8) are defined therein. The fluid Equations (1.8) coupled with Maxwell's Equations (1.3), where the charge and current densities are given by Equations (A.7), constitute a complete, closed, and self-consistent fluid description of a plasma.

1.3 The Ionosphere

The earth's entire atmosphere, starting from the surface and going up to space, can be subdivided into different layers that are characterized by their respective properties. In order of increasing altitude, these layers are referred to as the troposphere, stratosphere, mesosphere, and thermosphere. The majority of the

atmosphere (by mass) is contained in the troposphere, with the density (and pressure) drastically decreasing in altitude afterward. The temperature profile is more complicated and its various regions of rising and falling temperatures essentially characterize the different layers [2]. That being said, the breakdown of the atmosphere into these layers is not the end of the story, for it misses important details, such as the ozone layer within the stratosphere and more notably the *ionosphere* that coexists with the mesosphere and thermosphere.

1.3.1 Structure and Composition

The ionosphere is the upper region of the earth's atmosphere that has been partially ionized by the sun's radiation. It is the interface between the familiar neutral atmosphere and the space environment. The ionosphere is sufficiently ionized that it cannot be described by the equations of neutral fluid dynamics alone; the governing equations of plasma physics become essential and often the dominant players. That being said, the density of the neutral gas below 1000 km exceeds that of the ionospheric plasma, so the effects of neutrals particles cannot always be ignored. Much like the (neutral) atmosphere, the ionosphere too can be subdivided into layers characterized by their respective properties. In order of increasing altitude, these layers are the D region, E region, and F region. While there is no clear boundary for when the ionosphere ends, the topside F region around 1000 km is often considered to be this boundary. Some authors prefer the addition of another layer above the F region, referred to as the protonosphere. As its name suggests the protonosphere is largely filled with protons (and some helium ions), and marks the

end of the ionosphere. Above the ionosphere is the start of the *plasmasphere*, which extends to about 4-6 earth radii until the start of the *magnetosphere*.

The different layers of the ionosphere are essentially characterized by the ion composition, plasma density[1], and the wavelength of energetic solar radiation most efficiently absorbed by the neutral particles in that region. The latter is in fact the generation mechanism for the ionosphere: neutral atoms/molecules are ionized after interacting with high energy photons in a process called *photoionization*. The resulting mix of electrons, various ions and neutrals is what comprises the partially ionized plasma in the ionosphere. The D region plasma, produced by hard x-ray photoionization, is the lowest in altitude at approximately 70 - 90 km and is consequently accompanied by the highest neutral particle density in the ionosphere. The E region, at an altitude range of approximately 90 - 120 km, is produced by soft x-ray photoionization. The F region, with a rough altitude range of 120 - 1000 km, contains the densest plasma of the ionosphere that typically peaks around 250 or 300 km. The plasma here is created by extreme ultra-violet radiation, which is the lowest energy radiation that can result in ionization but is simultaneously more plentiful than soft/hard x-rays. The basic features of the neutral atmosphere and ionosphere described above are illustrated by Figure 1.1. Note that for the daytime ionosphere there is a subtle bump between the labels for the F region and E region. This small bump belongs to the F region and, when present, is referred to as the F_1 region. Consequently, the more familiar region near the density peak is referred to

[1]Recall that the total ion density (n_i) has to be very close to the density of electrons (n_e), by *quasineutrality* mentioned in the beginning: $n_i = \sum_s n_s \approx n_e$, where the sum is over the number density of the s^{th} ion species. Incidentally, this conveniently allows us to define $n \equiv n_e \approx n_i$ and talk about the *plasma density*, n, rather than n_i and n_e individually.

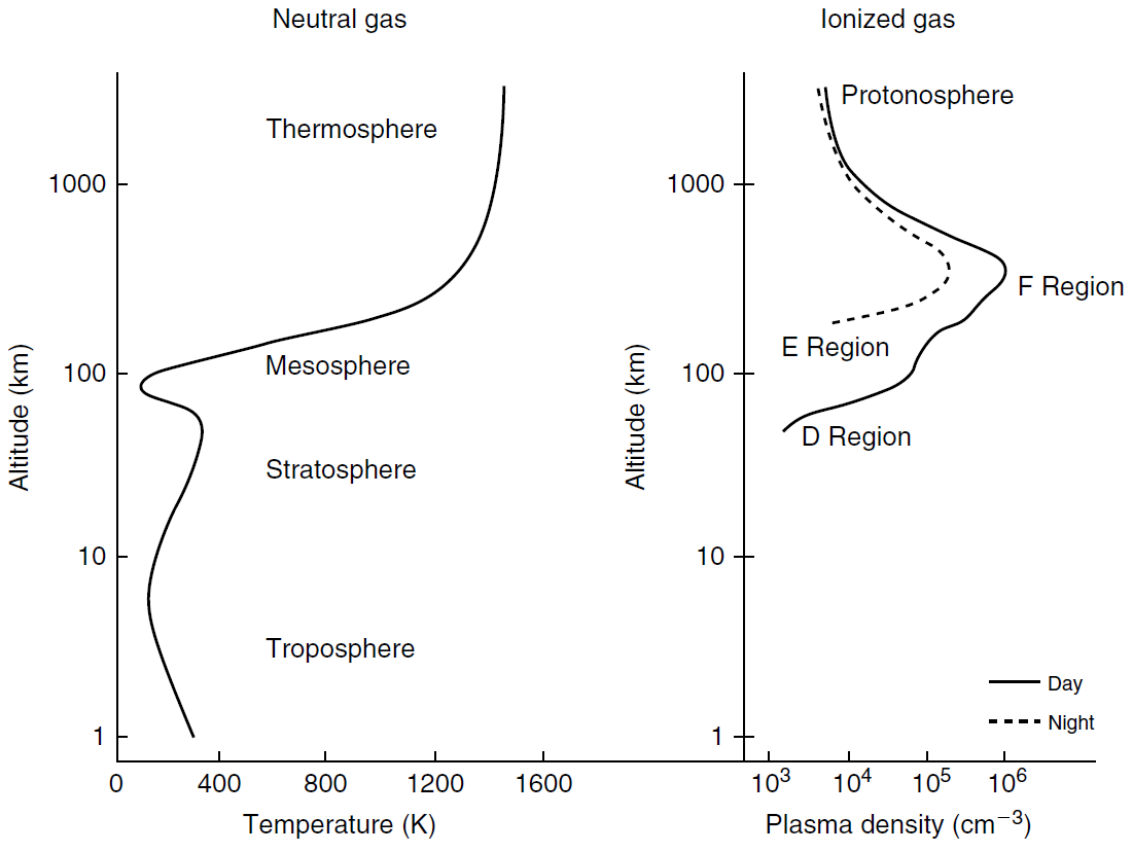

Figure 1.1: Illustration of the typical midlatitude profiles for the neutral atmosphere temperature and ionospheric plasma density as a function of altitude, with the various layers labeled. (Reproduced from [25].)

as the F_2 region; and the location of peak plasma density is termed the *F_2-peak*.

The neutral composition of the D and E regions are dominated by O_2 and N_2, while the F region is dominated by O and N_2. Atomic oxygen (O) is absent in the lower atmosphere since oxygen atoms like to pair up and make diatomic oxygen, implying that there must be a mechanism for their production at higher altitudes. This mechanism goes roughly as follows: solar UV radiation breaks up the O_2 molecules into two O atoms in a process called *photodissociation*, after which the O atoms being lighter than the surrounding gas rise upward and populate the F region via molecular diffusion [25]. This is reflected in the ion composition of the F region, which is mostly dominated by O^+, while the E region is predominantly O_2^+

and NO^+. The D region ion chemistry is rather different; the presence of several ions have been measured, such as O_2^+, NO^+, O_2^-, CO_3^-, HCO_3^-, NO_3^-, CO_4^-, among others [3]. Suffice it to say that the D region is quite complicated. Since the D region is peripheral to the dissertation, further discussion on its plasma chemistry will be omitted. During the nighttime, the sun's ionizing radiation is absent, and one would expect the ionosphere to disappear due to electron-ion *recombination*. As is suggested by Figure 1.1 this is not quite the case; it turns out that most of the ionosphere survives. This is because electrons and ions can recombine only if they lose energy through collisions with neutral particles, of which there are relatively few in the mid and upper ionosphere. At any rate, the losses are too minimal for significant recombination to take place before the sun rises again. The one caveat is that the rate of recombination of electrons and ions sharply depends on the ion species involved. It turns out that *dissociative* electron-ion recombination involving molecular ions (e.g. O_2^+ and NO^+) is much faster than *radiative* recombination of atomic ions (e.g. O^+) [25]. The meaning of these different types of recombination can be illustrated by the three examples below, where the first two are dissociative recombination while the last is radiative recombination:

$$O_2^+ + e^- \rightarrow O + O; \quad NO^+ + e^- \rightarrow N + O; \quad O^+ + e^- \rightarrow O + \gamma,$$

where γ is a photon typically in the UV range. With that in mind, it is no surprise that the D and E regions experience a rapid loss of plasma after dusk. The D region, being composed of only molecular ions and being the region of the ionosphere

with the highest neutral density, effectively disappears during the nighttime[2] due to recombination (see Figure 1.1). The E region survives during the nighttime but with low densities more characteristic of the (daytime) D region, while the F region being primarily composed of O^+ stays mostly intact. Altitudes above the F region are essentially unaffected. The ion and neutral composition for the altitude range 90 - 1000 km is shown in Figure 1.2, and for the reader's convenience Table 1.1

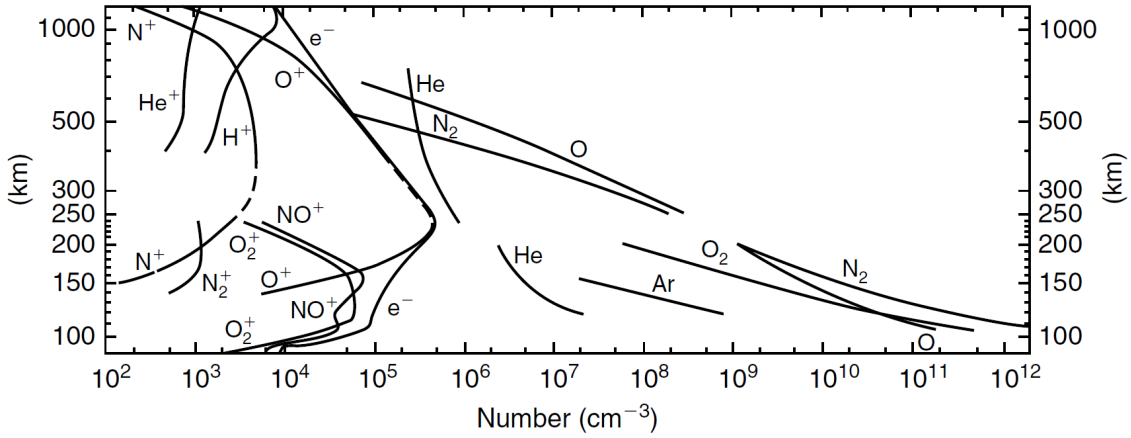

Figure 1.2: Daytime atmospheric composition based on mass spectrometer measurements above White Sands, New Mexico. The helium curve is measured during nighttime; curves above 250 km are from satellites. (Reproduced from [25].)

lists a number of important plasma parameters for various relevant altitudes. For a detailed discussion of the topics in this section, such as atmospheric/ionospheric structure and composition, plasma chemistry, experimental observations, theory, and much more, see the classic texts of Banks and Kockarts [2, 3].

Above the ionosphere is the start of the plasmasphere: the dense part of the plasma making up the earth's magnetosphere that can extend out to 6 earth radii in quiet times and be eroded back to 4 earth radii in active times; it is essentially

[2]Actually, this is somewhat of an oversimplification. It is true that without the radiation from the sun, the D region plasma density *drastically* plummets. But there is another source of ionization: the constant stream of background cosmic rays. This becomes the dominant ionization source during nighttime, albeit a much weaker one, that gives the D region a minimal (but non-zero) amount of plasma.

an extension of the ionosphere, composed of cold (eV) H^+ and He^+ ions. Above the plasmasphere is the remainder of the magnetosphere, a region of near-earth space filled with a tenuous plasma composed of protons and helium ions and electrons. The source of this plasma is the solar wind - a constant stream of supersonic plasma emanating from the sun. Unlike the plasmasphere, the solar wind plasma in the magnetosphere is hot (keV), and the helium is fully ionized (i.e. He^{2+}). As the solar wind impinges on the front of the magnetosphere it slows to subsonic speeds, creating a sharp change in solar wind density called the bow shock; it subsequently flows around and drags on the earth's magnetic field in the flow direction, thus giving the magnetosphere an elongated shape. In addition to plasma particles, the solar wind carries the sun's magnetic field. A property of highly conductive plasmas, called the *frozen-in theorem*, forces magnetic fields to be fixed in the plasma from which they originated; thus, the solar plasma drags on the sun's magnetic field and vice versa. An example of this is magnetic field lines emanating from the sun's surface and carrying the sun's plasma along, as is illustrated by the left panel of Figure 1.3 showing "plasma loops" near the sun's surface. Likewise, the outwardly flowing

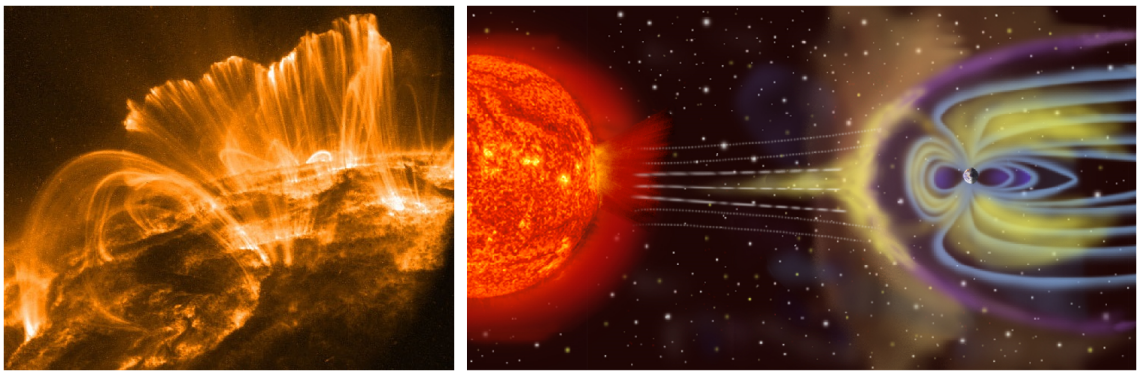

Figure 1.3: Photograph of the sun, and the earth's magnetosphere.

solar wind drags the sun's magnetic field with it, eventually coming in contact with

the earth's. When the sun's magnetic field is oriented opposite to the earth's, they can merge together and temporarily create open field lines that subsequently get dragged back and reconnect on the opposite side. This mechanism is called *magnetic reconnection* and is responsible for the dumping of vast amounts of high energy particles into the magnetosphere. In addition, the buffeting of the magnetosphere causes existing trapped electrons to precipitate down to the earth's atmosphere near the poles and crash into the neutral atoms, subsequently exciting them to release visible light of various colors that constitute the aurora borealis and australis. The right panel of Figure 1.3 shows an artist's rendition of the solar wind impinging on the earth's magnetosphere, creating the bow shock and the elongated structure of the magnetosphere toward the right of the figure.

1.3.2 Conductivity Tensor

The previous discussion on the structure and properties of the various ionospheric regions made it quite clear that the ionosphere is a partially ionized plasma. The last crucial property that has not been considered so far is the presence of earth's geomagnetic field $\mathbf{B}_0$ in the ionosphere. A plasma immersed in a sufficiently strong background magnetic field is said to be *magnetized*. Together with its partially ionized nature, the magnetized ionospheric plasma gives rise to a unique conductive response in the presence of electric fields. In fact the ionosphere acts as an anisotropic conductor at low frequencies, which we can show by considering the

linearized, steady state, cold[3] momentum equation for plasma species α:

$$0 = \frac{q_\alpha}{m_\alpha}(\mathbf{E} + \mathbf{u}_\alpha \times \mathbf{B}_0) - \nu_{\alpha n}\mathbf{u}_\alpha, \tag{1.9}$$

where $\nu_{\alpha n}$ is the α-neutral collision frequency. By defining $\Gamma_{\alpha n} \equiv \nu_{\alpha n}/\omega_{c\alpha}$, $\omega_{c\alpha} \equiv q_\alpha B_0/m_\alpha$, the parallel and perpendicular (to $\mathbf{B}_0$) components can be written as

$$u_\alpha^\parallel = \frac{1}{\Gamma_{\alpha n}}\frac{E_\parallel}{B_0}; \quad \Gamma_{\alpha n}\mathbf{u}_\alpha^\perp - \mathbf{u}_\alpha^\perp \times \hat{\mathbf{b}}_0 = \frac{\mathbf{E}_\perp}{B_0}, \tag{1.10}$$

where $\hat{\mathbf{b}}_0 = \mathbf{B}_0/B_0$ is the unit vector in the direction of the background field. Taking the cross product of the perpendicular equation in (1.10) with $\hat{\mathbf{b}}_0$ and using it to plug back in for $\mathbf{u}_\alpha^\perp \times \hat{\mathbf{b}}_0$, we can explicitly solve for $\mathbf{u}_\alpha^\perp$:

$$\mathbf{u}_\alpha^\perp = \frac{1}{B_0}\left[\frac{\Gamma_{\alpha n}}{1+\Gamma_{\alpha n}^2}\mathbf{E}_\perp + \frac{1}{1+\Gamma_{\alpha n}^2}\mathbf{E}_\perp \times \hat{\mathbf{b}}_0\right]. \tag{1.11}$$

We can find the current response due to $\mathbf{E}$ by substituting this result and $u_\alpha^\parallel$ into the expression for current density $\mathbf{J} = \sum_\alpha q_\alpha n_\alpha(\mathbf{u}_\alpha^\perp + u_\alpha^\parallel \hat{\mathbf{b}}_0)$. Moreover, taking $\hat{\mathbf{b}}_0 = \hat{\mathbf{z}}$ we can write the current density as $\mathbf{J} = \overline{\overline{\boldsymbol{\sigma}}} \cdot \mathbf{E}$, where $\overline{\overline{\boldsymbol{\sigma}}}$ is the conductivity tensor:

$$\mathbf{J} = \sigma_P \mathbf{E}_\perp - \sigma_H \mathbf{E}_\perp \times \hat{\mathbf{z}} + \sigma_\parallel E_\parallel \hat{\mathbf{z}} = \overline{\overline{\boldsymbol{\sigma}}} \cdot \mathbf{E}; \quad \overline{\overline{\boldsymbol{\sigma}}} \equiv \begin{pmatrix} \sigma_P & -\sigma_H & 0 \\ \sigma_H & \sigma_P & 0 \\ 0 & 0 & \sigma_\parallel \end{pmatrix}, \tag{1.12}$$

[3]Here "cold" refers to the neglect of thermal effects due to the pressure gradient term. This is the *cold-plasma approximation* and often valid in the ionosphere (more on this in Chapter 3).

where σ_P, σ_H, and $\sigma_\parallel$ are respectively the Pedersen, Hall, and parallel conductivities given by

$$\sigma_P = \sum_\alpha \tilde{\sigma}_\alpha \frac{\Gamma_{\alpha n}}{1+\Gamma_{\alpha n}^2}; \quad \sigma_H = \sum_\alpha \frac{\tilde{\sigma}_\alpha}{1+\Gamma_{\alpha n}^2}; \quad \sigma_\parallel = \sum_\alpha \frac{\tilde{\sigma}_\alpha}{\Gamma_{\alpha n}}, \tag{1.13}$$

and $\tilde{\sigma}_\alpha \equiv q_\alpha n_\alpha / B_0$. The Pedersen and parallel conductivities are intuitively expected because they represent the plasma current response in the direction of the applied electric field. However, the Hall conductivity determines the current response perpendicular to the applied electric field and $\mathbf{B}_0$, since it emerges due to the $\mathbf{E} \times \mathbf{B}_0$ drift of the charged particles. If for simplicity we assume there is only one (singly ionized) ion species (say, O^+) then (for $\alpha = e, i$) we have $n_e = n_i \equiv n_0$, and $-\tilde{\sigma}_e = \tilde{\sigma}_i \equiv \tilde{\sigma}$. Using typical profiles for $\nu_{en}(z)$, $\nu_{in}(z)$ and $n_0(z)$ as a function of altitude z and typical values for ω_{ce} and ω_{ci} we obtain the altitude dependent profiles depicted by Figure 1.4. It can be seen from the $\Gamma_{in}(z)$ profile that the ions are mag-

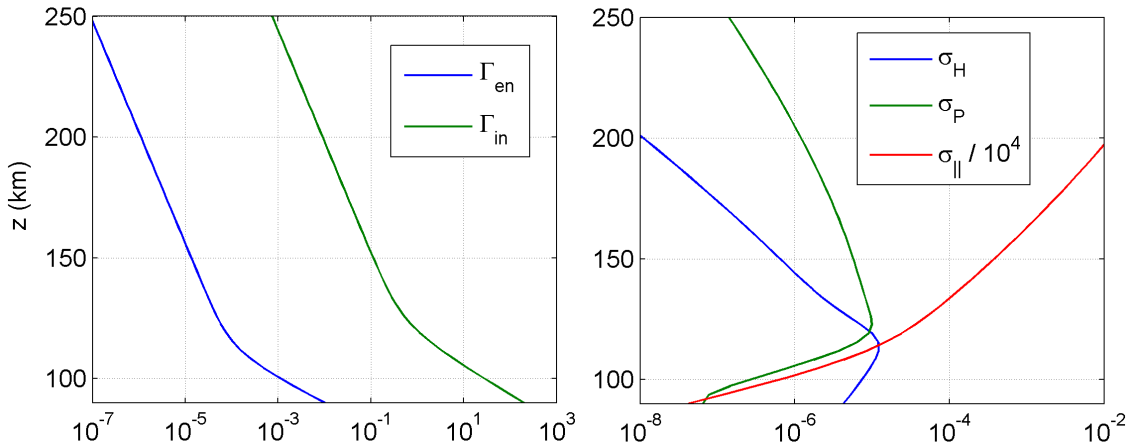

Figure 1.4: Ionospheric collision frequencies, conductivities, and plasma density.

netized for altitudes above 120 km ($\Gamma_{in} < 1$) but lose their magnetization ($\Gamma_{in} > 1$) below 120 km, which essentially signifies the start of the E region. For the domain

shown here, the electrons are magnetized ($\Gamma_{en}(z) < 1$) but become demagnetized in the bottom of the D region near 65 km. Note also that as the ions become demagnetized, the Hall conductivity rises rapidly. This is no coincidence: since ions make frequent collisions with neutrals, the electrons and ions can no longer $\mathbf{E} \times \mathbf{B}$ drift together in the presence of an electric field. Thus an overall current survives, and the Hall conductivity becomes significant. For this reason the E region is sometimes referred to as the Hall region. This Hall-dominant region is due to the presence of neutrals, and makes the ionospheric plasma unique since an analog of this does not exist in solar or astrophysical plasmas. The F region, like conventional plasmas, is Pedersen-dominant and supports the propagation of the usual cold-plasma wave modes since both the electrons and ions are magnetized. In the Hall region the ions are "glued" to the neutrals but electron motion survives, allowing only the propagation of waves supported by the electrons (e.g. whistler/helicon waves). The Hall region plays a major role for the topic discussed in Chapter 3.

1.4 Ionospheric Modification and Diagnostic Instruments

Since the beginning of its discovery, studying the ionosphere has been rather challenging since much of it is not directly accessible. That being said, ionospheric research took a huge leap forward due to the development of Low Earth Orbit (LEO) satellites and various ground based diagnostic instruments. While satellites have allowed the study of the mid to upper F region and magnetosphere, the lower altitudes have been less fortunate. The lower regions of the ionosphere are too high for balloon experiments and too low for satellites; the only direct access is with rockets exper-

iments, of which there have been relatively few attempts due to their considerable cost. One of the very first ground-based diagnostic instruments to be developed was the *ionosonde* - a kind of radar used for examining the ionosphere. The ionosonde transmits radio waves roughly in the 1 - 40 MHz range and has a receiver as well. The typical operation of the ionosonde constitutes sweeping the frequency range, waiting for the reflected signals to be received, and clocking the time of arrival of the reflected signal. Frequencies below the so-called *critical frequency* will be reflected back, while frequencies above it will pass through the ionosphere and never come back. The critical frequency is the maximum plasma frequency of the ionosphere and is given the label f_0F_2 while the altitude at which this occurs is termed h_mF_2. Of those frequencies that are reflected, the lower frequencies will arrive faster than the higher frequencies since they are reflected lower in the ionosphere. In fact, the reflection point of each frequency (f) occurs at the altitude (z) where it matches the local plasma frequency (f_p): $f = f_p(z)$ (more on this in Section 1.4.2). Moreover, $f_p(z)$ is proportional to the square root of the local plasma density; so keeping track of the time of arrival of the various frequencies sent up by the ionosonde gives a measurement of the height of reflection, while the frequency itself pins down the plasma frequency near the reflection height. Using this information it is possible to construct a profile of the plasma density vs. altitude, up to the F_2-peak. Altitudes above the F_2-peak are not accessible to the ionosonde, but simple theoretical models can be used to fit those altitudes instead.

Much of early ionospheric research was initially carried out by monitoring the ionosphere in its natural state, such as day/night/seasonal variations, effects of in-

creased solar activity, etc. While such studies are important and still active research areas today, there has since been the addition of *ionospheric modification* studies. Rather than waiting for ionospheric changes to be induced by the sun, scientists began inducing changes from the ground by using arrays of antennas capable of transmitting HF radio waves of considerable power. These waves propagate into the ionosphere and are partially absorbed near the reflection altitude, causing a considerable rise in electron temperature and a zoo of linear and nonlinear effects that are still actively being studied today. Some of these effects are[4]: Langmuir/upper-hybrid turbulence, field aligned plasma transport, plasma wave generation, parametric instabilities, and mode conversion. The typical goal of ionospheric modification experiments is no surprise: run the experiment, gather data from as many diagnostic sources as possible, and compare the results against theoretical models. In a sense, the ionosphere became an *open plasma laboratory*, unlike man-made laboratory plasmas that necessarily have *boundaries* that contain the plasma. Since the use of high-power HF waves proved to be so fruitful in studying the basic plasma physics of the ionosphere, several facilities were built around the world that housed arrays of antennas with this capability. These facilities can modify the ionosphere on demand, often with precise control of the frequency and amount of power being injected. Since the initial effect is to heat up the ionospheric plasma, these facilities came to be called *ionospheric heaters*. There are several ionospheric heaters around the world today, some still operational while others closed down. Figure 1.5 shows a map of the world with the various ionospheric heaters indicated. The HIPAS and

[4]This list is by no means exhaustive and does not represent the entirety of current ionospheric research, but rather reflects the author's knowledge bias and research interests.

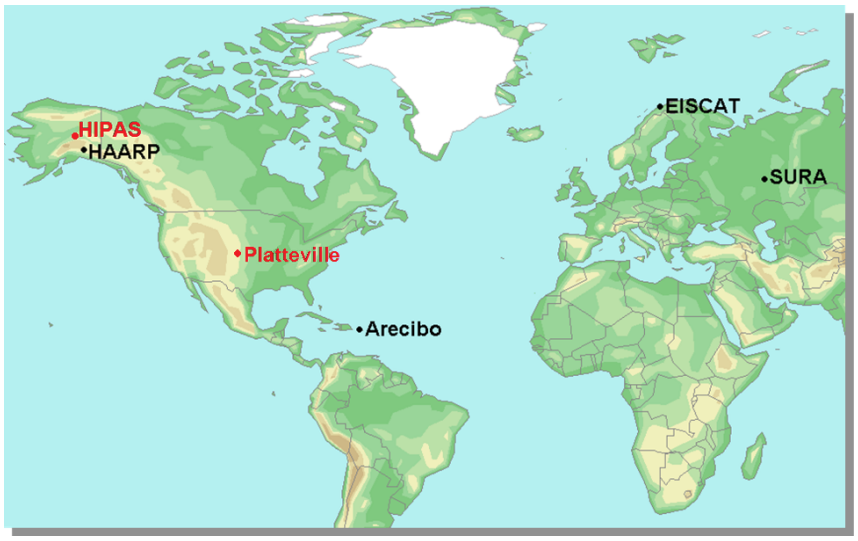

Figure 1.5: Various ionospheric heaters around the world. Those marker in red are no longer in service.

Platteville heaters (marked in red) are not in service today, Arecibo is coming online in the near future, and EISCAT and SURA are in operation. The High Frequency Active Auroral Research Program (HAARP) facility is the world's most powerful ionospheric heating facility and the most relevant for this dissertation; essentially all discussions pertain to results of experiments carried out at HAARP.

Heating facilities have various on-site diagnostics that help scientists set up their ionospheric heating experiments and measure subsequent effects. All have an on-site ionosonde, since it can gather quite useful information about the state of the ionosphere and its rough density profile, while being relatively simple and inexpensive. This allows scientists to tweak the parameters of the HF transmitter to be most effective for the experiment at hand. For example, conducting an experiment that involves the peak density of the F region will require an HF frequency that is near the peak plasma frequency, which is easily obtained by an ionosonde. Another important instrument is the *Incoherent Scattering Radar* (ISR) - an indispensable tool used for ionospheric studies first proposed by Bill Gordon in 1958. An ISR uses

radar dishes that can operate at very/ultra high frequencies and sends these radio waves into the ionospheric plasma. These waves then scatter off the random fluctuations in the plasma and are subsequently received on the ground and processed. Amazingly, analysis of these scattered signals can give quantitative measurements of the electron density, electron temperatures, ion temperature, ion composition and plasma velocity. Moreover, it can measure all of these quantities well above the F_2-peak, unlike the ionosonde. The caveat is that an ISR is *much* more expensive than an ionosonde; currently, only the EISCAT and Arecibo facilities have an ISR.

1.4.1 The HAARP Facility

The HAARP facility is the world's foremost ionospheric research facility, located near Gakona, Alaska. Since its beginning in 1990, HAARP has made several important contributions to ionospheric research due to its high power and frequency flexibility. The cornerstone of HAARP is the *ionospheric research instrument*, a high-power HF radio frequency transmitter capable of sending up to 3.6 MW of power into the ionosphere. Equally impressive is a phased array of 12×15 cross-dipole antennas. The antenna array can be phased to generate a narrow radio beam pointed in various different directions, in certain situations reaching a maximum Effective Radiated Power (ERP) of 5 GW. The HF frequency range is quite flexible, allowing anything from 2 - 10 MHz with a wave polarization of one's choosing (typically left/right hand circular polarization). It is even possible to generate quick frequency hops or smoothly frequency chirp during an experiment; and subdivide the array into two and control each section separately. All of this put together

has allowed scientists to perform several novel experiments over the years. Figure

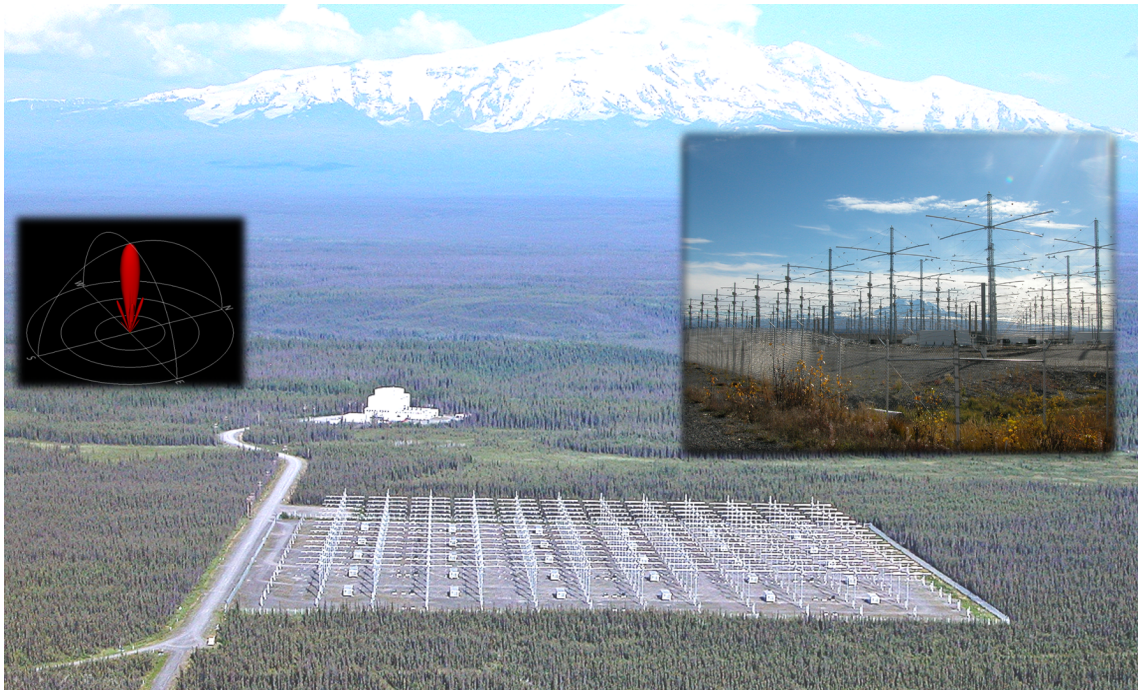

Figure 1.6: The HAARP facility photographed from above; right insert is a close-up of the array; left insert is a computer generated beam pattern.

1.6 shows a photograph of the HAARP facility from above. The 12 × 15 antenna array is clearly seen, along with the control center in the back. The right insert shows a close-up of the cross-dipole array. The left insert is a typical HF beam pattern, generated on a computer by a model of the HAARP antenna array. There are several on-site diagnostic instruments at HAARP, such as magnetometers, optical photometers, HF/VHF radars, ULF/ELF/VLF receivers, ionosonde, and more. In addition there are the overflying satellites DEMETER, DMSP, CLUSTER, and others. Much of the data to be presented in subsequent chapters will focus on data collected by the LEO satellites DEMETER[5] and DMSP. The ionosonde at HAARP is often used to check the ionospheric condition prior to doing an experiment. Con-

[5]DEMETER actually went offline December 2010, so it is no longer taking data. Nevertheless, much valuable data was collected while it was still active during HAARP experiments that is still being analyzed today.

sistent and "good" reflections of the HF radar waves that are picked up by the ionosonde paint a picture like the one depicted in Figure 1.7. Such figures are called *ionograms* and include radar wave reflections (shown as colored squares), computed plasma frequency profile as a function of altitude (solid black line), and estimates of several ionospheric parameters (not shown), such as the critical frequency, critical altitude, and many more.

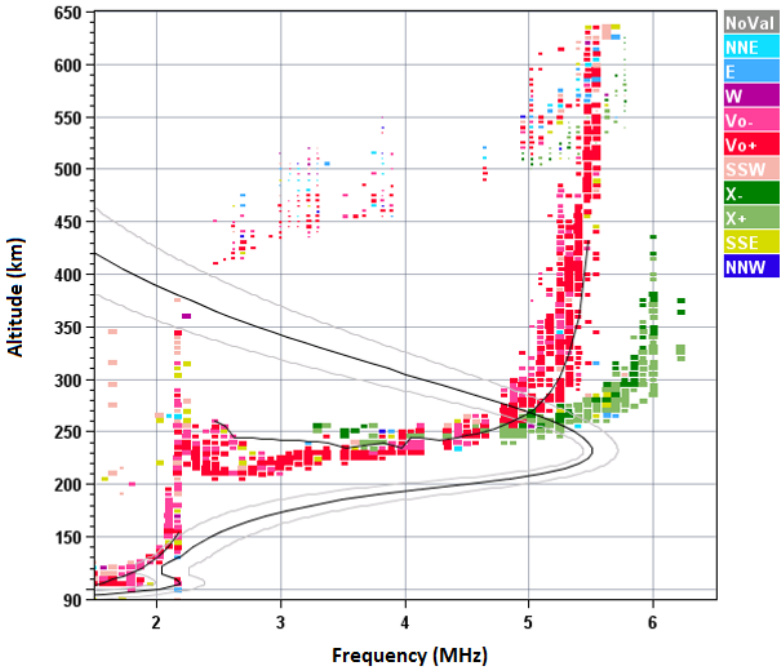

Figure 1.7: Typical ionogram produced by the ionosonde showing various reflected radar waves (red and green squares) and computed plasma frequency profile at the current time (solid black curve), along with those for before/after (solid gray curves).

1.4.2 Physics of Ionospheric Modification

As soon as the HF waves leave the radiation zone of the antenna array, their propagation within the atmosphere is straightforward and to large extent analytically tractable. However, when the HF waves reach the ionosphere, their propagation is significantly complicated by the fact that the ionospheric plasma acts as a

nonuniform, anisotropic dielectric medium. The medium being *nonuniform* refers to the fact that the index of refraction is space dependent, which in the context of the ionosphere primarily refers to altitude dependence; *anisotropic* refers to the index of refraction being direction dependent. There are three important directions to keep in mind for wave propagation in plasmas, and those are the directions of the wavevector ($\mathbf{k}$), wave electric field ($\mathbf{E}$), and the background geomagnetic field ($\mathbf{B_0}$). How these vectors are oriented relative to each other (among other things) determines the speed of propagation of the HF radio wave. The theory of radio wave propagation in a cold magnetized plasma was worked out by several scientists, including E. V. Appleton, D. Hartree, H. K. Lassen, and K. G. Budden. The primary result, which is the index of refraction, can be obtained by starting with the linearized fluid/Maxwell system of equations and using the so-called *cold plasma approximation* (more on this in Section 3.1.1). The expression for the index of refraction (η) in a cold magnetized plasma (see [9] and Appendix B.1) is given by the *Appelton-Hartree equation*:

$$\eta^2 = 1 - \frac{X}{U - \frac{1}{2}\frac{Y^2(1-(\hat{\mathbf{k}}\cdot\hat{\mathbf{b}})^2)}{(U-X)} \pm \sqrt{\left(\frac{1}{2}\frac{Y^2(1-(\hat{\mathbf{k}}\cdot\hat{\mathbf{b}})^2)}{U-X}\right)^2 + Y^2(\hat{\mathbf{k}}\cdot\hat{\mathbf{b}})^2}}, \tag{1.14}$$

where the dot products are between the unit wavevector, $\hat{\mathbf{k}} = \mathbf{k}/k$, and the unit vector in the direction of the background field, $\hat{\mathbf{b}} = \mathbf{B_0}/B_0$. Equivalently the dot product can be expressed in terms of the angle (θ) between the wavevector and background magnetic field, that is $\hat{\mathbf{k}}\cdot\hat{\mathbf{b}} = \cos(\theta)$. The variables included are defined as follows:

$$X = \frac{\omega_p^2}{\omega^2}, \quad \omega_p^2 = \frac{n_e e^2}{m_e \epsilon_0}, \quad Y = \frac{\omega_c}{\omega}, \quad \omega_c = \frac{|eB_0|}{m_e}, \quad Z = \frac{\nu}{\omega}, \quad \text{and } U = 1 - iZ,$$

where ω_p, ω_c, and ν are the electron plasma, gyro, and collision frequencies, respectively. The two solutions associated with the $\pm$ are referred to as the *ordinary* (O) and *extraordinary* (X) modes, respectively. When an ionospheric heater is operating with O-mode it means the HF waves being emitted from the array are Left Hand Circularly Polarized (LHCP), while X-mode[6] means the HF waves are Right Hand Circularly Polarized (RHCP)[7]. The ray path for the HF waves can be computed from Eq. (1.14), which is essentially the *dispersion relation* of the waves since $\eta = ck/\omega$. The simplest case is when magnetization and collisional effects can be neglected, that is, when $Y \to 0$ and $Z \to 0$. In this case the index of refraction is

$$\eta^2 = 1 - X = 1 - \frac{\omega_p^2}{\omega^2}. \tag{1.15}$$

Or to put it another way:

$$\omega^2 = \omega_p^2 + c^2 k^2, \tag{1.16}$$

which is true for both the O-mode and X-mode since magnetization is neglected. Light rays that propagate up will be reflected at or below the altitude where the wave experiences a *cutoff*. A cutoff occurs when the index of refraction vanishes ($\eta = 0$), which in this case implies $1 - X = 0$, or $\omega = \omega_p(z_r)$, where z_r is the reflection altitude.

[6]This terminology can be rather confusion for someone starting out in this field because the standard nomenclature from cold plasma wave theory has the L-, R-, O-, and X-waves, which are *special cases* of Eq. (1.14). To make things worse the O/X-waves *are not* left/right hand circularly polarized (but the L/R-waves are, as their names suggest). On the bright side, for parallel propagation ($\theta = 0$), the O/X-modes are consistent with the L/R-waves; likewise, for perpendicular propagation ($\theta = \pi/2$), the O/X-modes are consistent with the O/X-waves.

[7]Wave polarization in plasma physics is typically defined relative to $\mathbf{B_0}$. A RHCP wave is one whose electric field rotates in the "right hand sense" when the thumb of the right hand is pointing in the direction of $\mathbf{B_0}$ (opposite is true for LHCP). These definitions are not universal, optical physics has a definition of circular polarization relative to the *wavevector* $\mathbf{k}$ instead of $\mathbf{B_0}$.

The ray path for this simple case is straightforward and leads to two different pictures depending on the wave frequency in relation to the critical frequency (f_c). For wave frequencies above f_c, the waves pass through the ionosphere, while frequencies below f_c are reflected at their respective cutoff altitudes. Figure 1.8 shows typical ray paths for these two scenarios, computed using the simplified index of refraction (1.15) and a typical ionospheric density profile with $f_c = 2$ MHz. All rays in Figure 1.8a are 2.1

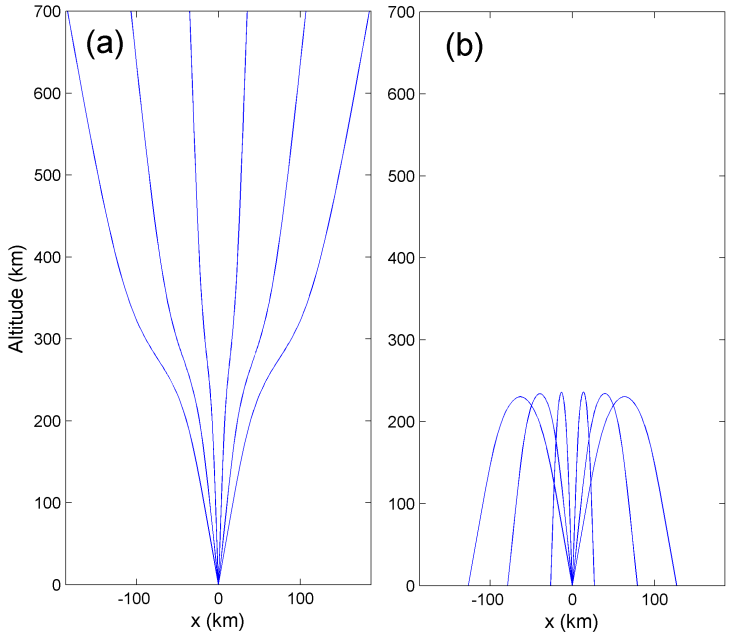

Figure 1.8: Typical ray paths for HF wave frequencies above (a) and below (b) the critical frequency.

MHz while those in Figure 1.8b are 1.9 MHz. Those that pass through nevertheless "feel" the effect of the ionospheric plasma, which is why rays bend (refract) near the altitude of maximum plasma density.

The above treatment with Equation (1.15) is very simplified; so much so that the difference between the O/X-modes has vanished. However, in practice the physics of ionospheric modification is quite different for each case, and to have a clearer understanding of the differences, we must go back to the Appleton-Hartree

equation (1.14). Let us assume collisional effects can be neglected ($Z = 0$) and the HF waves are launched very close to vertically, so that θ is small (but $\theta \neq 0$). With these assumptions we can neglect the first term under the square root of Equation 1.14 and are left with the quasi-longitudinal approximation:

$$\eta^2 = 1 - \frac{X}{1 - \frac{1}{2}\frac{Y^2 \sin\theta^2}{1-X} \pm Y\cos\theta}. \tag{1.17}$$

Letting $\sin\theta \to \theta$ and $\cos\theta \to 1$ for simplicity, and setting $\eta = 0$ gives the cutoff condition for the O/X-modes:

$$1 - \frac{1}{2}\frac{Y^2\theta^2}{1-X} \pm Y - X = 0. \tag{1.18}$$

Multiplying through by $(1 - X)$ gives a quadratic equation in X, whose solution is given by

$$X = \frac{1}{2}\left(2 \pm Y + (-1)^n Y\sqrt{1 + 2\theta^2}\right) \approx 1 \pm Y\left(\frac{1 \mp (-1)^n}{2}\right) + \frac{1}{2}\theta^2 Y, \tag{1.19}$$

where $n = 0, 1$ correspond to the two possible solutions of the quadratic. Now letting $\theta \to 0$ and choosing[8] $n = 0$ the cutoff condition simplifies to

$$X = 1 \pm Y\left(\frac{1 \mp 1}{2}\right). \tag{1.20}$$

[8]Finding the proper reflection altitude for the O/X modes is actually somewhat subtle; the $n = 1$ solution is not chosen since it corresponds to the so-called "Z-mode", which is not of interest for the studies presented in this dissertation.

Thus for the O-mode we have $X = 1$, while for the X-mode we have $X = 1 - Y$. The reflection altitude of the O-mode (z_O) is the same as before: $\omega = \omega_p(z_O)$. We can find a similar expression for the X-mode by assuming $Y < 1$ (which is essentially always true for F region ionospheric heating) and expanding:

$$\frac{\omega_p}{\omega} = \sqrt{1 - Y} \approx 1 - \frac{1}{2}\frac{\omega_c}{\omega}. \tag{1.21}$$

The reflection altitude for the X-mode (z_X) is then (approximately) given by $\omega = \omega_p(z_X) + \frac{1}{2}\omega_c$. Note that if we had set $\theta = 0$ from the beginning in Equation (1.18), then we would have lost the $X - 1 = 0$ solution entirely and instead had the solution $X = 1 \pm Y$ for the O/X-modes. Evidently there is some kind of "jump" in the reflection altitude of the O-mode when going from $\theta \neq 0$ to $\theta = 0$ [16, Ch.5]. However, this is not of practical importance since there is always some nonzero angle that will allow the wave to "turn around" and reflect the usual way.

It is clear from comparing $\omega = \omega_p(z_X) + \frac{1}{2}\omega_c$ and $\omega = \omega_p(z_O)$ that the X-mode reflects at a lower altitude than the O-mode, giving the O-mode access to a larger altitude range. Near the reflection point of both the O-mode and X-mode there is a "swelling" of the HF pump[9] electric field [28]. This is illustrated for O-mode in Figure 1.9 taken from Figure 3b of [28]. The component of the pump electric field parallel to the geomagnetic field (tall solid curve, marked as $E_\parallel$) experiences an increase as the wave approaches the reflection point (marked by z_O), where it takes on an Airy pattern as one would expect. The two perpendicular components

[9]The term *pump* is often used in plasma physics in reference to the incident wave, since it is the source of power that "pumps" energy into the system. The terms *HF wave*, *HF pump*, *pump wave*, or simply *pump*, will be used interchangeably throughout.

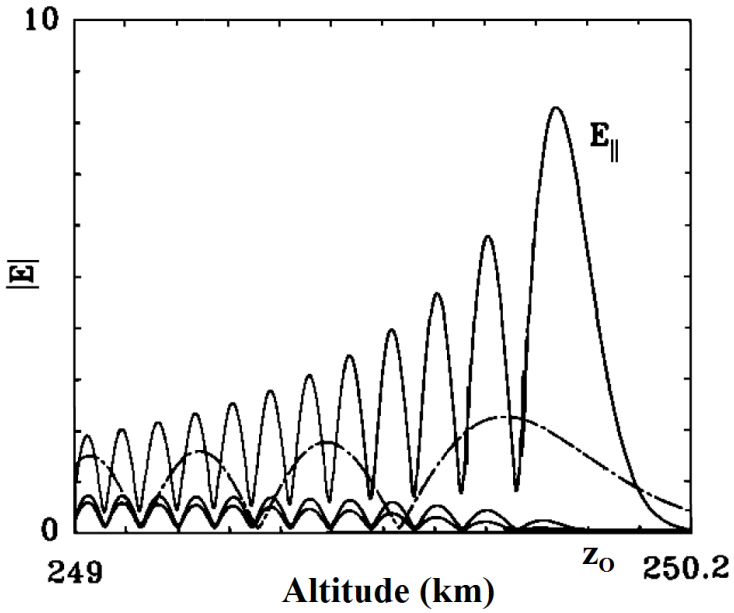

Figure 1.9: Electric field of an O-mode wave as a function of height near the reflection point z_O of a linear electron density profile.The solid curve labeled $E_{||}$ represents the parallel component, while the other two solid curves are the perpendicular components. For comparison, the field obtained for an unmagnetized plasma is given by the dot-dashed curve. The vertical scale is volts per meter if the up-going wave is normalized to 1 V/m at 100 km height. (Adapted from [28, Fig. 3b].)

(short solid curves) decrease as the wave transforms from an ElectroMagnetic (EM) wave to an electrostatic wave. For comparison the electric field pattern is shown for propagation in an unmagnetized plasma (dot-dashed curve). The substantial increase in the longitudinal (parallel) component of the pump electric field drives electrostatic plasma waves (Langmuir waves). Moreover, the pump electric field is sufficiently high for various nonlinear processes to take place in this region, such as parametric instabilities, ponderomotive forces, and ultimately *Langmuir turbulence*.

Another region of the ionosphere that experiences instabilities is below the reflection point by about 3 - 10 km. This region is near the point when the wave frequency equals the so-called Upper-Hybrid (UH) frequency - the frequency (ω_{uh}) associated with plasma oscillations perpendicular to the background field, given by $\omega_{uh}^2 = \omega_p^2 + \omega_c^2$. The Lorentz force from the background field gives an additional

restoring force to the electrons that they otherwise would not have, thus raising their frequency of oscillation from ω_p to $\sqrt{\omega_p^2 + \omega_c^2}$. The pump frequency matches the UH frequency at the *upper-hybrid altitude*, z_{uh}, given by the condition $\omega^2 = \omega_p^2(z_{uh}) + \omega_c^2$. If we again assume $\omega_c < \omega_p$ then we can approximate this condition as $\omega \approx \omega_p(z_{uh}) + \frac{1}{2}\omega_c \left(\frac{\omega_c}{\omega_p}\right)$. Comparing with the reflection altitude conditions of the O/X modes we see that the UH altitude is always below z_O but always above z_X:

$$z_O > z_{uh} > z_X.$$

Thus only the O-mode has access to the UH altitude, making the X-mode of little interest for most of the topics discussed throughout. A resonance is established when an O-mode pump wave passes through the UH altitude, called Upper-Hybrid Resonance (UHR), that converts a significant portion of the pump wave energy into the bulk plasma and raises its temperature considerably. The turbulence that ensues is referred to as *upper-hybrid turbulence*, and is the unifying theme for most of the topics covered in the dissertation.

Alt. (km)	n_e (10^5 cm^{-3})	f_{pe} (MHz)	$n(O^+)$ (%)	$n(O_2^+)$ (%)	$n(NO^+)$ (%)	T_e (K)	T_i (K)	T_n (K)	f_{ce} (MHz)	f_{ci} (Hz)	f_{lh} (kHz)	f_{uh} (MHz)	r_{ce} (cm)	r_{ci} (m)	λ_{De} (cm)	λ_e (m)	v_{te} (10^5 m/s)	v_s (10^3 m/s)
120	0.52	2.05	0.00	0.42	0.58	345	345	345	1.50	26.59	6.32	2.55	0.77	1.82	0.56	23.22	0.72	0.30
135	0.62	2.24	0.02	0.46	0.52	543	489	489	1.49	26.70	6.32	2.69	0.97	2.16	0.64	21.30	0.91	0.38
150	0.72	2.41	0.09	0.46	0.45	740	579	579	1.48	28.28	6.48	2.83	1.14	2.26	0.70	19.76	1.06	0.45
165	0.92	2.72	0.24	0.36	0.39	938	635	635	1.47	31.63	6.83	3.09	1.29	2.21	0.70	17.55	1.19	0.53
180	1.56	3.55	0.39	0.26	0.34	1136	671	671	1.46	34.93	7.15	3.84	1.43	2.15	0.59	13.43	1.31	0.62
195	2.40	4.40	0.54	0.16	0.29	1333	693	693	1.45	38.58	7.49	4.64	1.56	2.07	0.51	10.84	1.42	0.70
210	3.22	5.10	0.66	0.11	0.23	1521	730	707	1.44	41.16	7.71	5.30	1.67	2.07	0.47	9.36	1.52	0.77
225	3.71	5.47	0.77	0.08	0.15	1685	777	716	1.43	43.32	7.88	5.65	1.77	2.11	0.46	8.72	1.60	0.85
240	3.72	5.48	0.85	0.05	0.09	1840	824	722	1.42	44.73	7.98	5.66	1.87	2.19	0.49	8.71	1.67	0.92
255	3.58	5.38	0.92	0.03	0.05	1995	871	725	1.41	46.53	8.11	5.56	1.96	2.21	0.51	8.88	1.74	0.98
270	3.37	5.21	0.96	0.01	0.02	2150	918	728	1.40	46.65	8.09	5.40	2.05	2.33	0.55	9.16	1.81	1.04
285	3.10	5.00	0.98	0.01	0.01	2303	965	729	1.40	47.03	8.10	5.19	2.13	2.37	0.59	9.54	1.87	1.08
300	2.82	4.77	1.00	0.00	0.00	2431	1012	730	1.39	47.17	8.08	4.97	2.20	2.44	0.64	10.00	1.92	1.12
315	2.54	4.53	0.98	0.00	0.00	2487	1059	731	1.38	48.08	8.13	4.73	2.24	2.45	0.68	10.54	1.94	1.14
350	1.95	3.96	0.98	0.00	0.00	2557	1169	731	1.36	48.22	8.08	4.19	2.31	2.57	0.79	12.04	1.97	1.15
400	1.32	3.27	0.97	0.00	0.00	2653	1326	732	1.33	47.94	7.97	3.52	2.41	2.76	0.98	14.61	2.01	1.18
450	0.92	2.72	0.97	0.00	0.00	2750	1480	732	1.30	48.37	7.92	3.01	2.51	2.89	1.20	17.56	2.04	1.20
500	0.65	2.30	0.96	0.00	0.00	2847	1629	732	1.27	50.19	7.98	2.62	2.61	2.93	1.44	20.77	2.08	1.22
550	0.48	1.97	0.94	0.00	0.00	2942	1779	732	1.24	53.19	8.12	2.33	2.71	2.90	1.70	24.17	2.11	1.25
600	0.37	1.72	0.93	0.00	0.00	3029	1928	732	1.21	59.15	8.48	2.11	2.81	2.74	1.98	27.70	2.14	1.27
650	0.29	1.52	0.91	0.00	0.00	3115	2077	732	1.19	66.66	8.90	1.93	2.91	2.54	2.27	31.32	2.17	1.30
700	0.23	1.36	0.90	0.00	0.00	3201	2227	732	1.16	71.55	9.12	1.79	3.01	2.46	2.57	34.98	2.20	1.33
800	0.16	1.13	0.88	0.00	0.00	3373	2525	732	1.12	76.00	9.21	1.59	3.23	2.49	3.19	42.36	2.26	1.37
850	0.13	1.04	0.87	0.00	0.00	3455	2675	732	1.09	78.63	9.27	1.51	3.33	2.49	3.51	46.05	2.29	1.40
900	0.11	0.96	0.86	0.00	0.00	3514	2824	732	1.07	81.87	9.36	1.44	3.43	2.47	3.83	49.70	2.31	1.42
1000	0.09	0.84	0.83	0.00	0.00	3493	3123	732	1.03	88.89	9.56	1.33	3.56	2.42	4.37	56.91	2.30	1.43

Table 1.1: A list of several ionospheric plasma parameters at various altitudes. From left to right the columns are: altitude, electron density, and plasma frequency; O^+, O_2^+, and NO^+ concentrations; electron, ion, and neutral temperatures; electron and ion gyrofrequencies; lower and upper hybrid frequencies; electron and ion gyroradii; electron Debye and inertial lengths; electron thermal speed and ion acoustic (sound) speed.

Chapter 2

Effects of Ionospheric Heating: Plasma Transport

Section 1.4 discussed how ionospheric heating facilities can be used to drastically raise the temperature of the ionospheric plasma. However, heating of the ionospheric plasma (say, in the F region) has an additional effect. Namely, there is an accompanying decrease in plasma density due to the tendency of a hotter region of plasma (like an ordinary gas) to expand outward away from the heated region and thus leave behind a density depression. Note that it is the increase in *electron* pressure and temperature that drives the electrons to expand initially. Since ions are much more massive and efficiently transfer energy to neutrals via collisions, their temperature changes very little and so they initially stay where they are. As the electrons try to escape from the heated region (while the ions stay) an electric field, known as the ambipolar electric field, is set up due to charge separation that drags the ions along. So while only the electrons are heated, the plasma as a whole expands out of the modified region. Since charged particles experience the Lorentz force perpendicular to $\mathbf{B_0}$ but are allowed to free stream parallel to $\mathbf{B_0}$, the expanding plasma travels primarily along the background field in the upward and downward directions. This movement of plasma along field lines is referred to as

field-aligned plasma transport and will be the main topic of this chapter.

2.1 Basic Physics of Ionospheric Ducts

Localized heating of the ionospheric plasma results in a increased pressure that expels plasma from the modified region, driving it along the geomagnetic field. So while the plasma near the modified region has a density depletion, the portion of the plasma that is transported upward constitutes an overall positive field-aligned density enhancement that stretches well into the topside F region and magnetosphere. Since the transverse density gradients of these structures are "bell-shaped" , they are essentially artificially generated "plasma tubes" and are thus referred to as *artificial ionospheric ducts*[1]. While artificial ducts are our primary concern, it is worth mentioning that they are one of many possible field-aligned density structures. In addition there are *natural ducts* and *natural negative ducts* (troughs), where the former constitutes a positive density enhancement of natural (rather than artificial) origin, while the latter is an inverted version of the former, i.e. a field-aligned density depletion. Lastly, it is possible to have naturally occurring field-aligned density structures whose transverse gradients involve single-density slopes rather than a bell shape. The presence of any of these field-aligned density structures plays a critical role in the propagation of whistler waves in the ionosphere and magnetosphere. The density structures serve as wave guides for VLF/ELF waves since the density gradient perpendicular to the magnetic field can lead to their total internal reflection [53], similar to light propagation in fibers. Such density structures have been observed

[1]This is somewhat cumbersome to rewrite all the time, so it will be used interchangeably with *artificial ducts*, *ionospheric ducts*, or simply *ducts*. Whether they are artificial in origin or not will be clear from the context.

[10] to extend over distances covering entire magnetic field lines. They are known to trap and guide whistler waves between conjugate regions of the geomagnetic field (e.g. [26]). For strong density gradients that extend along the entire field line, VLF waves injected at a magnetic footpoint will continuously experience internal reflect until they are transmitted to the conjugate footpoint, as depicted by Figure 2.1 [42]. In addition, whistler waves generated from lightning sources and VLF transmitters

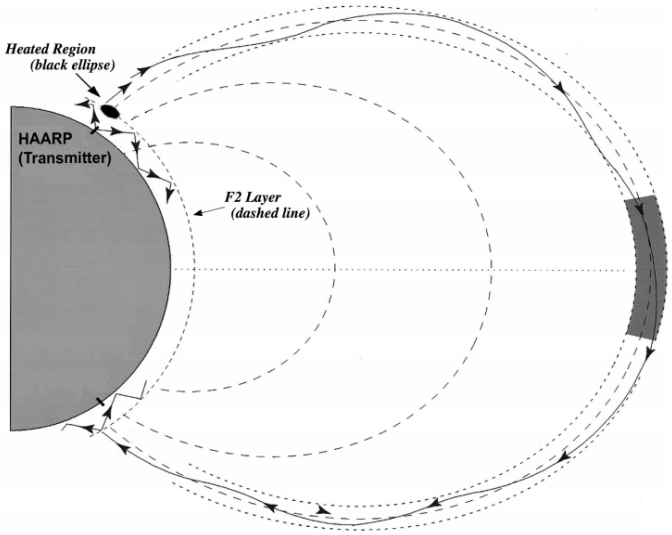

Figure 2.1: Schematic illustrating how an ionospheric heating-induced duct that stretches along an entire field line can guide VLF waves from one magnetic footpoint to its conjugate. (Reproduced from Reference [42].)

propagate along the geomagnetic field $\mathbf{B_0}$ into the magnetosphere and thus play an important role in the radiation belts dynamics (e.g. [26, 10]). Naturally occurring ionospheric ducts are common due to solar wind activity or otherwise, but the main focus here will be ionospheric ducts artificially generated by the HAARP ionospheric heater.

The possibility for artificially creating trans-hemispheric ducts was discussed by Reference [42], where the SAMI2 ionosphere model [24] was used to simulate the artificial heating-induced plasma structure along the magnetic dipole flux tube.

It was shown that long term continuous HF heating of the F region by powerful ionospheric heaters, such as HAARP, generates a strong thermal wave in the ionospheric and magnetospheric plasma. The thermal wave propagates along the magnetic field line through the topside ionosphere and magnetosphere, driving ion outflows, displacing the ambient plasma and leading to the formation of density ducts that stretch along the magnetic field line to the conjugate point.

A qualitative picture of artificial duct formation and the physics involved can be demonstrated by a crude model of the plasma transport effects during ionospheric heating. The model is based on the fluid momentum equations for a partially ionized magnetized plasma with electrons and one ion species, given by

$$mn\left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{u}\right) = -\boldsymbol{\nabla}p \pm en(\mathbf{E} + \mathbf{u}\times\mathbf{B}) - mn\nu\mathbf{u}.$$

Here the $\pm$ indicates the sign of the charge, ν is the effective electron collision frequency, and subscripts indicating the species have been omitted for simplicity. As was mentioned above, the plasma transport is primarily along the field lines, which we take to be oriented in the $\hat{\mathbf{z}}$ direction. So, we can approximate the flow velocity as $\mathbf{u} \simeq u_z\hat{\mathbf{z}}$ and only consider the parallel component of the momentum equation:

$$mn\left(\frac{\partial u_z}{\partial t} + u_z\frac{\partial u_z}{\partial z}\right) = -\frac{\partial p}{\partial z} \pm enE_z - mn\nu u_z.$$

Assuming a steady state ($\partial u_z/\partial t = 0$) and that u_z is sufficiently small (or ν is sufficiently large), we can ignore the convective derivative [12, Ch.5] on the LHS

and solve for u_z to give

$$u_z = \pm\mu E_z - D\frac{\partial_z n}{n} - \frac{1}{m\nu}\frac{\partial T}{\partial z},$$

where $\mu = e/m\nu$ is the *mobility* and $D = T/m\nu$ is the *diffusion coefficient*. The flux of each species, $\mathbf{\Gamma}_\alpha$, also only has a parallel component: $\Gamma_\alpha\hat{\mathbf{z}} = nu_{z,\alpha}\hat{\mathbf{z}}$ $(\alpha = e, i)$, where $n = n_e = n_i$ is the plasma density (required by quasineutrality). Noting that only the electron temperature changes considerably during ionospheric heating $(\partial T_i/\partial z \approx 0)$, we can write the electron and ions fluxes individually as

$$\Gamma_{z,e} = -\mu_e n E_z - D_e\frac{\partial n}{\partial z} - \frac{n}{m_e\nu_e}\frac{\partial T_e}{\partial z}, \tag{2.1a}$$

$$\Gamma_{z,i} = \mu_i n E_z - D_i\frac{\partial n}{\partial z}. \tag{2.1b}$$

Since the electron flux is larger than the ion flux due to electron heating (driven by the external HF heating source $\partial T_e/\partial z$), the electrons will leave behind a positive charge and set up an (ambipolar) electric field with a polarity that will drag the ions along. So the ions and electrons move together, which mathematically translates to $\Gamma_{z,e} = \Gamma_{z,i} \equiv \Gamma_z$. Using this requirement to solve for E_z, and plugging it back into Equation (2.1)a (or (2.1)b) gives

$$\Gamma_z = -D_a\frac{\partial n}{\partial z} - \frac{\mu_i n}{m_e\nu_e(\mu_e + \mu_i)}\frac{\partial T_e}{\partial z}, \tag{2.2}$$

where the new effective coefficient D_a is referred to as the *ambipolar diffusion coef-*

ficient [12, Ch.5] and is given by

$$D_a = \frac{\mu_i D_e + \mu_e D_i}{\mu_i + \mu_e}.$$

The continuity equation (A.12) tells us that

$$\frac{\partial n}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{\Gamma} = 0, \tag{2.3}$$

but this is not quite accurate for the ionosphere since there is a constant ionization source that controls the background ionospheric density profile. Sources of ionization, labeled below as S_0 (and assumed to be static in time), can be added to the RHS of (2.3):

$$\frac{\partial n}{\partial t} + \boldsymbol{\nabla} \cdot (\Gamma_z \hat{\mathbf{z}}) = S_0. \tag{2.4}$$

Using Equations (2.2) and (2.4), assuming that $\mu_e >> \mu_i$ [12, Ch.5], and D_a approximately constant[2], we find that the plasma density is given by[3]

$$\frac{\partial n}{\partial t} = D_a \frac{\partial^2 n}{\partial z^2} + \frac{n}{m_i \nu_i} \frac{\partial^2 T_e}{\partial z^2} + S_0. \tag{2.5}$$

This is a *diffusion equation* with diffusion coefficient D_a and a dynamic source term proportional to $\partial_z^2 T_e$. It is convenient to break up the quantities into the ambient and perturbed (due to heating) parts, so that $n = n_0 + \delta n$ and $T_e = T_0 + \delta T_e$.

[2]This is not quite true, at least in the heated region, since D_a has a dependence on T_e; but that is not the point since the end goal is a qualitative understanding.

[3]There should be another term here proportional to $\frac{\partial n}{\partial z}\frac{\partial T_e}{\partial z}$, which is an advection term. It was neglected since it has no effect outside of the heated region (where $\partial T_e/\partial z = 0$); our main concern is with terms that can generate plasma transport well outside the heated region.

Finally, noting that S_0 must by definition satisfy

$$\frac{\partial n_0}{\partial t} = D_a \frac{\partial^2 n_0}{\partial z^2} + \frac{n_0}{m_i \nu_i} \frac{\partial^2 T_0}{\partial z^2} + S_0,$$

Equation (2.5) reduces to

$$\frac{\partial \delta n}{\partial t} = D_a \frac{\partial^2 \delta n}{\partial z^2} + \frac{n_0}{m_i \nu_i} \frac{\partial^2 \delta T_e}{\partial z^2}. \tag{2.6}$$

For simplicity weak heating has been assumed, so that $\delta n/n_0$ is small; terms that are of order $\delta n/n_0$ smaller compared to others have thus been neglected. Finally, normalizing the variables in Equation (2.6) gives

$$\frac{\partial \bar{n}}{\partial \bar{t}} = \frac{\partial^2 \bar{n}}{\partial \bar{z}^2} + C \frac{\partial^2 \bar{T}_e}{\partial \bar{z}^2}. \tag{2.7}$$

where $\bar{n} = \delta n/n_0$, $\bar{T}_e = \delta T_e/T_{max}$, $C = T_{max}/(m_i \nu_i D_a)$, T_{max} is the maximum temperature perturbation attained by δT_e, $\bar{z} = z/L$, and $\bar{t} = t/\tau$, where $\tau = L^2/D_a$ is the *diffusion time* for length-scale L. Since the HF pump wave tends to be absorbed in a narrow region of the F region ionosphere, we can assume a localized temperature distribution. For simplicity, let $\bar{T}_e$ be given by a Gaussian with a half-width of unity that turns on at $\bar{t} = 0$:

$$\bar{T}_e = e^{-\bar{z}^2} H(\bar{t}), \tag{2.8}$$

where $H(.)$ is the Heaviside step function. Defining the Fourier transform pairs

$$\tilde{f}(k) = \int_{-\infty}^{\infty} f(z)\ e^{-ikz}\ dz \equiv \mathcal{F}(f), \tag{2.9a}$$

$$f(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(k)\ e^{ikz}\ dk \equiv \mathcal{F}^{-1}(\tilde{f}), \tag{2.9b}$$

and Fourier transforming Equation (2.7) (with Equation (2.8) plugged in) gives

$$\frac{\partial\, \tilde{n}(k,t)}{\partial\, t} + k^2 \tilde{n}(k,t) = -k^2 C \mathcal{F}(e^{-z^2}) H(t) = -k^2 C \sqrt{\pi} e^{-\frac{1}{4}k^2} H(t), \tag{2.10}$$

where the "bar" over all variables have been temporarily dropped for simplicity. Initially the system is undisturbed, which mathematically translates to $\tilde{n}(k,0) = 0$. With this initial condition, the solution for $\tilde{n}(k,t)$ can be readily obtained:

$$\tilde{n}(k,t) = C\sqrt{\pi} e^{-\frac{1}{4}k^2} \left(1 - e^{-k^2 t}\right) H(t). \tag{2.11}$$

Taking the inverse Fourier transform takes us back to real space: $n(z,t) = \mathcal{F}^{-1}(\tilde{n})$. Carrying out the integration and putting the "bars" back in we obtain the full solution:

$$\bar{n}(\bar{z},\bar{t}) = C \left(\frac{e^{-\frac{\bar{z}^2}{4\bar{t}+1}}}{\sqrt{4\bar{t}+1}} - e^{-\bar{z}^2} \right) H(\bar{t}). \tag{2.12}$$

Letting $C = 0.1$ and plotting $\bar{n}(\bar{z},\bar{t})$ for several values of $\bar{t}$ shows the evolution of the density after heating is turned on, as shown by Figure 2.2. At $t = 0$ the density perturbation is completely zero; soon afterward it becomes negative near $\bar{z} = 0$, while further out there is a density perturbation enhancement. A negative $\delta n/n_0$ implies a drop in density relative to the background, and it occurs exactly where the

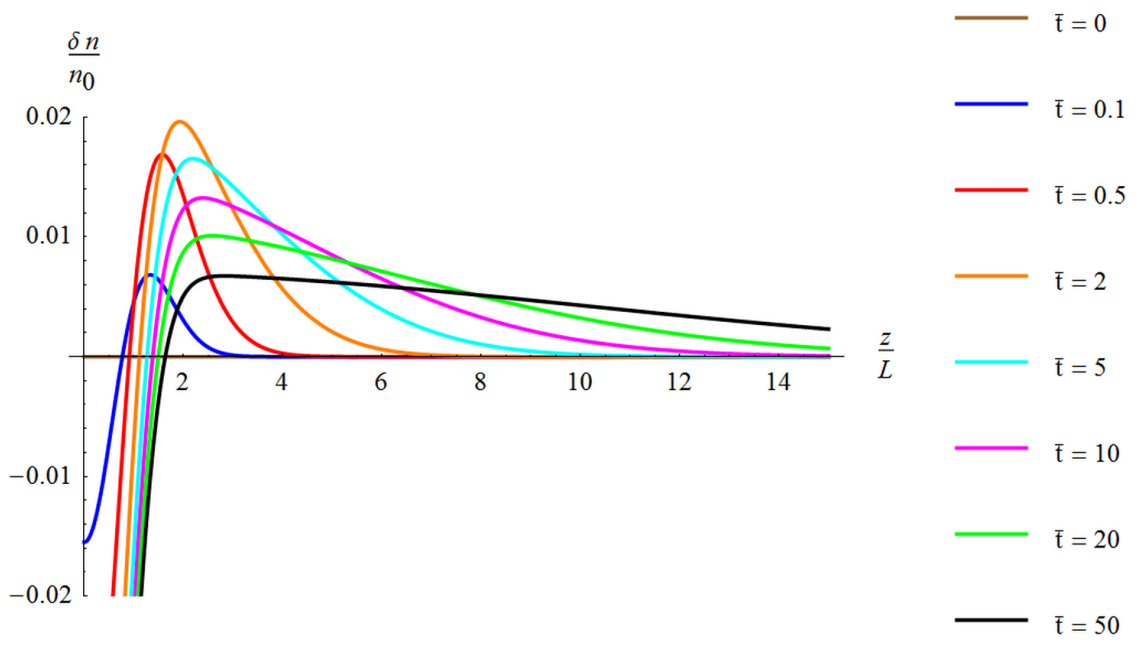

Figure 2.2: Ambipolar diffusion of relative plasma density perturbation after pump turn on; legend indicates different values of $\bar{t}$.

heating source is non-zero; such behavior was alluded to earlier on in the chapter.

This gives a rough qualitative picture of what happens during ionospheric heating and how plasma from the heated region can populate the field lines far away, thus creating artificial ionospheric ducts. That being said, this very crude model misses (among others) three very important aspects of the physics. The first is the lack of a self-consistent ionization source; the one used here is static and was assumed to have no dependence on the density itself. A more realistic model should have a first principles approach to calculating the rate of ionization. Secondly, the model assumes the electron/ion inertia terms are negligible. While this is accurate for the electrons due to their small mass, this is not necessarily true for the ions. Disregarding ion inertia yields the diffusion equation (2.7), which does not obey causality. Including ion inertia will give a finite speed of propagation for the perturbed density "pulse" generated by ionospheric heating. Finally, many of

the simplifying arguments used to make things tractable can only be valid *locally*. A comprehensive description of ionospheric duct formation requires a *global* model valid for a large range of altitudes and latitudes.

2.2 Model of Ionospheric Duct Formation

To properly describe the physics of artificial ionospheric duct formation, a computational model of the ionosphere was adopted and modified for the purposes of our study. Here we present this theoretical model of ionospheric duct formation due to HF heating that will be used throughout this chapter. The model is based on a modification of the SAMI2 [24] code, and is validated by comparison with a well documented experiment. The HF heating experiment was conducted at the EISCAT HF heating facility and diagnosed by the EISCAT ISR. The ISR produced real-time profiles of the electron and ion temperature between 150 - 600 km, and are shown in Section 2.2.3 to be in good agreement with the model.

2.2.1 The SAMI2 Model of the Ionosphere

The theoretical/computational model is based on the SAMI2 model developed at the Naval Research Laboratory [24]. SAMI2 treats the dynamic plasma and chemical evolution of seven ion species (H^+, He^+, N^+, O^+, N_2^+, NO^+, and O_2^+) in the altitude range of about 100 km to several thousands of kilometers. The model includes $\mathbf{E} \times \mathbf{B}$ drift of the field lines with frozen-in plasma (in altitude and longitude), an empirical neutral atmosphere model, horizontal winds, photo-deposition into the ionosphere, ion chemistry models, and ion inertia. This inclusion

of ion inertia is critical since it allows for the study of sound wave propagation in the plasma. SAMI2 is a global, inter-hemispheric model and can simulate the plasma along the entire dipole magnetic field line (for the geometry of the model see Figure 2.1 or [42]).

The basic equations of SAMI2 are the fluid transport equations, which are worked out in Section A.2. Due to differences in notation and form of the SAMI2 model equations in Reference [24], they are quoted below for completeness; the ion continuity, momentum, and temperature equations for each ion species i are given by[4]

$$\frac{\partial n_i}{\partial t} + \boldsymbol{\nabla} \cdot (n_i \mathbf{V}_i) = \mathcal{P}_i - \mathcal{L}_i n_i, \tag{2.13a}$$

$$\frac{\partial \mathbf{V}_i}{\partial t} + \mathbf{V}_i \cdot \boldsymbol{\nabla} \mathbf{V}_i = -\frac{1}{m_i n_i} \boldsymbol{\nabla} P_i + \frac{e}{m_i} \mathbf{E} + \frac{e}{m_i c} \mathbf{V}_i \times \mathbf{B} + \mathbf{g} - \nu_{in} (\mathbf{V}_i - \mathbf{V}_n) - \sum_{j \neq i} \nu_{ij} (\mathbf{V}_i - \mathbf{V}_j), \tag{2.13b}$$

$$\frac{\partial T_i}{\partial t} + \mathbf{V}_i \cdot \boldsymbol{\nabla} T_i + \frac{2}{3} T_i \boldsymbol{\nabla} \cdot \mathbf{V}_i - \frac{2}{3} \frac{1}{n_i k} \boldsymbol{\nabla} \cdot (\kappa_i \boldsymbol{\nabla} T_i) = Q_{in} + Q_{ii} + Q_{ie}, \tag{2.13c}$$

where n_i is the ion density, $\mathbf{V}_i$ is the ion flow velocity, $\mathcal{P}_i$ represents the ion production terms, and $\mathcal{L}_i$ represents the ion loss terms[5]; in Equation (2.13b) P_i is the ion pressure, $\mathbf{B}$ is the external geomagnetic field, $\mathbf{g}$ is the gravitational field[6], ν_{in} is the ion-neutral collision frequency, $\mathbf{V}_n$ is the neutral velocity, and ν_{ij} is the ion-ion

[4]These equations are written in Cartesian coordinates. However, equations in SAMI2 are solved using dipole coordinates (since a dipole magnetic field is assumed), but for brevity that will not be reproduced here.

[5]These terms involve photoionization, radiative recombination, and chemistry; they are discussed in detail in [24]. Note that unlike the continuity equation (A.12) derived in Appendix A, Equation (2.13a) has source terms due to ionization/recombination processes in the ionosphere.

[6]Note that essentially the only difference between Equation (2.13b) and the momentum equation (A.17) from Appendix A is the addition of the gravitational field, which was neglected in the latter for simplicity.

collision frequency; κ_i is the ion thermal conductivity and the heating terms (or loss terms, depending how you think about it) on the RHS of (2.13c)[7] are due to ion-neutral collisions (Q_{in}), ion-ion collisions (Q_{ii}), and ion-electron collisions (Q_{ei}). Equation (2.13b) is solved for every ion species in the model, but Equation (2.13c) is solved for three ion species: H^+, He^+, O^+; the temperatures of the molecular ions N_2^+, NO^+, and O_2^+ is set equal to the O^+ temperature.

The electron density is found by quasineutrality, while the electron momentum and temperature equations in SAMI2 are given by

$$0 = -\frac{1}{m_e n_e}\boldsymbol{\nabla} P_e - \frac{e}{m_e}\mathbf{E} - \frac{e}{m_e c}\mathbf{V}_e \times \mathbf{B}, \tag{2.14a}$$

$$\frac{\partial T_e}{\partial t} - \frac{2}{3}\frac{1}{n_e k} b_s^2 \frac{\partial}{\partial s}\kappa_e \frac{\partial T_e}{\partial s} = Q_{en} + Q_{ei} + Q_{phe}. \tag{2.14b}$$

In Equation (2.14a) electron inertia has been neglected courtesy of the small electron mass. Electron collision terms have been neglected as well because $\nu_e << \Omega_e$, where ν_e denotes the electron collision frequency and Ω_e is the electron gyrofrequency. The electron temperature equation (2.14b) is written in dipole coordinates; s is along the dipole field line, k is Boltzmann's constant, b_s is the magnitude of Earth's magnetic field normalized to the equatorial field on the ground, Q_{en} is heating due to electron-neutral collisions, Q_{ei} is due to electron-ion collisions, Q_{phe} is due to photoelectron heating, κ_e is the electron thermal conductivity, and only parallel thermal conduction has been assumed.

During discretization of the above transport equations (once they are changed

[7]Note that Equation (2.13c) is just like the energy equation (A.18) from Appendix A, but with the loss terms redefined on the RHS.

to dipole coordinates), SAMI2 considers only their parallel (field-aligned) components. Thus, with the exception of the $\mathbf{E} \times \mathbf{B}$ drift, there is *no cross-field transport* in SAMI2. For further details and definitions of quantities (e.g. collision frequencies and heating terms) not reproduced here, see [24] and references therein.

2.2.2 Addition of Localized Heating Source

Since the SAMI2 model does not consider wave propagation and absorption, we introduced in the model a flexible source of electron heating due to HF wave absorption, as was done by Reference [42] and subsequent papers on the topic. This HF electron heating source, represented by Q_{HF}, was added to the RHS of the SAMI2 temperature equation:

$$\frac{\partial T_e}{\partial t} - \frac{2}{3}\frac{1}{n_e k} b_s^2 \frac{\partial}{\partial s}\left(\kappa_e \frac{\partial T_e}{\partial s}\right) = Q_{en} + Q_{ei} + Q_{phe} + Q_{HF},$$

We chose Q_{HF} in the form of a localized heating rate per electron:

$$Q_{HF} = Q_0 \; e^{-(z-z_0)^2/\Delta z^2} e^{-(x-x_0)^2/\Delta x^2}, \tag{2.15}$$

where Q_0 is the peak electron heating rate (in K/s) and is typically given by

$$Q_0 = \frac{\mu P}{k_B n_e V}. \tag{2.16}$$

Here P is the power of the HF heater, n_e is the local electron density, μ is the absorption efficiency ($0 < \mu < 1$), V is the volume of the HF heated region, and Δz

and Δx are its vertical and North-South (N-S) horizontal half-widths. The vertical offset, z_0, is set close to the UH altitude where heating occurs (see Section 1.4.2), while x_0 represents the N-S horizontal offset of the heated region. The HF-irradiated spot is an ellipse with N-S angular half-width Θ and East-West (E-W) angular half-width Φ, with the horizontal and vertical half-widths given by $\Delta x = z_0 \tan(\Theta)$ and $\Delta z = z_0 \tan(\Phi)$, respectively; Θ and Φ depend on the irradiated frequency and have to be found from the engineering specifications of the HF transmitter. It is assumed that electron heating occurs in an altitude range having vertical extent Δz between the HF wave reflection point and the UH altitude, which is dominated by the anomalous absorption [19]. With these parameters the volume can be approximated as $V = \pi \Delta x \Delta y \Delta z$.

The procedure for simulating ionospheric heating is as follows: the code starts up from empirically determined initial conditions 24 hours before the specific heating time, and runs for 24 hours of "world clock time". This practice allows the system to relax to ambient conditions, and reduces noise in the system due to the initialization. Furthermore, the neutral density model is often adjusted so that the f_0F_2 computed by SAMI2 matches the value observed by the ionosonde during an experiment. Then the "artificial heater" turns on and begins to pump energy into the electrons, using the specified parameters for that run. Artificial heating continues for some time, continuously pumping energy into the electrons at the specified altitude, and the perturbations in ion and electron properties are tracked as they travel along the field line; the heater switches off after a specified time, allowing the ionosphere to relax back to ambient conditions. A typical ionospheric heating run performed using the

modified SAMI2 is illustrated by Figure 2.3, which shows a snapshot of the relative plasma density (n_e/n_0) after 10 minutes of heating at an altitude of approximately 300 km. Ionospheric duct formation in the form of plasma enhancements ($n_e/n_0 > 1$)

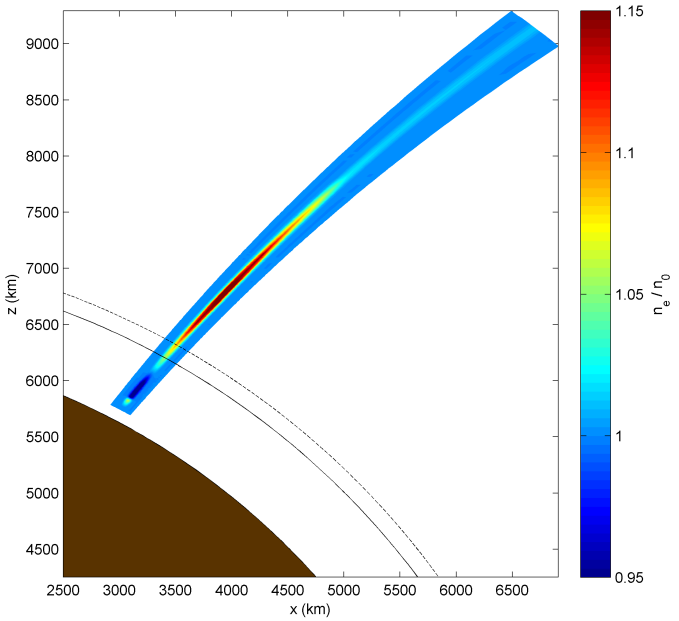

Figure 2.3: Typical ionospheric heating run using SAMI2, showing n_e/n_0 (colorbar), the earth (brown shape), and the DEMETER (solid line) and DMSP (dashed) orbits cutting through the ionospheric duct; the axes (in km) start at the center of the earth.

are seen above and below the heating source; at the location of the source there is a plasma depletion ($n_e/n_0 < 1$), as was predicted from the crude model in Section 2.1. The brown shape at the bottom left of the figure represents the earth, while the solid and dashes lines cutting through the simulation represent the orbits of the LEO satellites DEMETER and DMSP, with respective altitudes of about 700 and 850 km above the ground; the figure axes are measured in km from the center of the earth.

In order to isolate and measure the perturbations directly, one run without artificial heating is performed where Q_0 in (2.15) is set to zero. We refer to this as the "ambient" or "reference" run, while those with artificial heating are "heated"

runs. The ionosphere changes during a simulation due to natural causes (albeit not that much for our time scales), so the perturbations in the heated runs are mixed with these natural variations. But since the same natural variations are present in the ambient data, scaling (or subtracting) by the ambient data provides a simple way to decouple the natural variations from the heater induced perturbations. Figure 2.3 is in fact an example of this, since the "n_0" was obtained from an ambient run.

2.2.3 Benchmarking the Modified SAMI2

To benchmark the modified SAMI2 (i.e. the addition of the artificial heating source) we modeled a particular well-documented experiment during 10/07/99 at the time of an EISCAT experiment [47]. We therefore use in the SAMI2 code the corresponding A_p[8] and $F_{10.7}$[9] indexes, and assumed that the heating began 10/07/1999 at 19:24 Universal Time (UT). The radiated HF power was 960 kW, the half power beam width was 12°, and the facility was operated at a frequency of 4.5 MHz. Furthermore, for the unperturbed profile of the electron density we find that the reflection height for the 4.5 MHz frequency is located at 280 km, while the UH altitude is about 10 km below that. The vertical extent of the anomalous absorption region is therefore taken as $\Delta z = 10$ km.

For the specified heater and antenna characteristics at EISCAT, Equation (2.16) becomes $Q_0 = 12400\ \mu$ (K/s). In our runs the heating rate was varied in the range 2000 - 8000 K/s, corresponding to an absorption efficiency range of $\mu = 0.16$

[8]This index is a measure of the general level of geomagnetic activity over the globe for a given day.

[9]This is a measure of the noise level generated by the sun at a wavelength of 10.7 cm at the earth's orbit. This value is put into ionospheric models to take into account solar output in wavelengths that produce photoionization in the ionosphere.

- 0.64. Note that Reference [22] used the radar data collected during the heating experiments at Tromso to estimate the heating rate per electron as 3000 K/s. This value corresponds to the absorption efficiency $\mu = 0.25$, which is within the range of our estimates. SAMI2 simulation results of the electron density during ionospheric heating are shown in Figure 2.4 [31]. The altitude profile of the electron density is

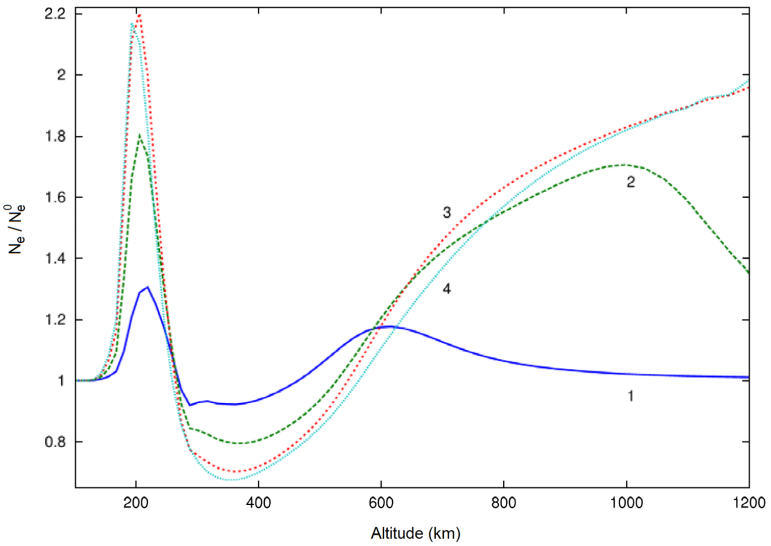

Figure 2.4: The computed normalized electron density at different times (numbered) for a pumping rate of 8000 K/s ($\mu = 0.64$). Traces 1 - 3 are separated by 3 minutes, where trace 1 is 1 min and 46 s into heating; trace 4 corresponds to cooling over 2 minutes and 49 seconds. (Reproduced from Reference [31].)

normalized to its ambient value and computed at different times for a given pumping rate $Q_0 = 8000$ K/s, which corresponds to an absorption efficiency $\mu = 0.64$. The heating was switched on at 19:24:00 UT for 8 minutes. The traces labeled 1 to 3 correspond to times separated by 3 minutes starting at 19:25:46, i.e. 1 min and 46 s into the heating. The trace 4 corresponds to cooling over 2 minutes and 49 seconds. The figure reveals that the electron heating increases the plasma pressure and thus pushes the plasma from the heated region along the magnetic field line. Consequently, the plasma density in the heated region drops, but on a timescale

larger than 5 minutes, as shown by the trace 3. Note that the crude ambipolar diffusion model results in Figure 2.2 are in qualitative agreement with the results seen here[10].

Figure 2.5 [31] shows the results of the model superimposed onto the observation results presented in Figure 3 of Reference [47], where the latter were obtained from the EISCAT ISR at 19:28 UT. The left panel shows the observed altitude profile of the electron density (circles) and that computed by the SAMI2 model (continuous trace) during 4 minutes into heating. The middle panel shows the observed ion temperature (circles) and electron temperature (crosses) along with three traces generated by SAM12 model. In order to improve agreement between the model and observations the neutral density in the model was adjusted so that the computed f_0F_2 peak matches the observations. For this purpose we have reduced the density of the atomic oxygen in the model by 50%. Such an approach is justified by the fact that SAMI2 uses long-term-averaged neutral density values in its atmosphere model, which need not be highly accurate for any specific day. The adjustment leads to significant changes in the electron temperature and affects the vertical velocity only slightly. The dash, solid, and dot-dash lines in Figure 2.5 correspond to the absorption efficiencies $\mu = 0.16$, 0.32 and 0.64, respectively. Note that the changes in μ affect only the values of electron temperature, while the ion temperature remains unperturbed during a relatively short heating pulse. The rightmost panel shows ion velocity observations (diamonds), along with three traces again corresponding

[10]Of course, you can only compare Figure 2.4 above the heating altitude since the results in Figure 2.2 assumed an infinite line to the left and right of the heating source, while for the realistic SAMI2 case there is a "wall" encountered by the plasma in the E region 100 km below the heating altitude, where the plasma "piles up" and then lost due to recombination.

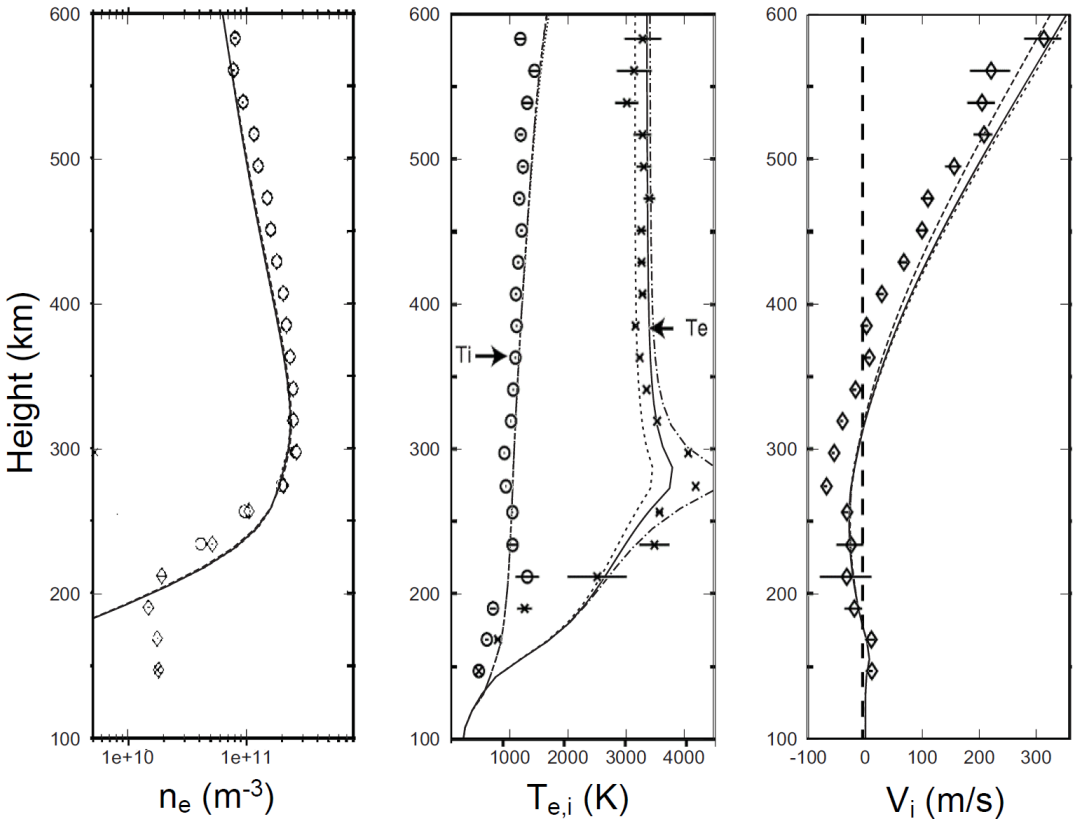

Figure 2.5: SAMI2 modeling results of ionospheric heating superimposed onto altitude (km) dependent data collected by the EISCAT ISR; (left) observed electron density (circles) and SAMI2 model computation (continuous trace) during 4 minutes into heating; (middle) observed ion temperature (circles) and electron temperature (crosses), along with three traces generated by the model; (right) observed ion velocity (diamonds) along with three traces, corresponding to computations made at different absorption efficiencies μ = 0.16, 0.32 and 0.64 (from left to right trace).

to the computations made for absorption efficiencies $\mu = 0.16$, 0.32 and 0.64 (from the left trace to right trace). For the HF heating duration considered here, the time scale of $\mathbf{E} \times \mathbf{B}$ drift was not exceeded; as such, the energy loss due to horizontal transport can be neglected.

Figure 2.5 reveals that HF heating with the absorption efficiencies $\mu = 0.3$ - 0.6 drives perturbations of the electron temperature in good agreement with those detected by the ISR. Moreover, the computed ion velocity fits well with the observations. Namely, it shows that the ion velocity is negative below the heating region,

and positive above it. Strong electron heating increases the electron pressure and pushes the plasma both down and upward from the heated region. Thus below this region the ion velocity is negative (downward directed); above the region it is positive (upward directed), and its value increases with altitude since the plasma density monotonically decreases.

2.3 Comparison of Model and Observations

Here we report on satellite observations of artificial ionospheric ducts in the topside ionosphere during HAARP campaigns in 2009 - 2010, funded by the BRIOCHE program[11]. The artificial ducts, generated by high-power HF radio wave injections from the HAARP transmitter toward the Magnetic Zenith (MZ), are detected by the DEMETER and DMSP satellites on a regular basis when there is a perceptible ionospheric F_2 peak density. Overall, the plasma density enhancements detected between 09:30 and 12:30 LT varied from 3 - 13%, while those during 17:30 - 22:15 LT were typically 15 - 40%. The modified SAMI2 model is used to study the artificial duct formation driven by HF electron heating in the bottomside F_2 region, similar to Section 2.2.3. The simulation results, performed for the input parameters similar to the conditions of the heating experiments, are in agreement with the pertinent observations. The ducts seem to be produced most efficiently for heating frequencies quite close to the critical frequency f_0F_2.

[11]Basic Research on IOnospheric CHaracteristics and Effects (BRIOCHE) was a DARPA program with the aim of exploring ionospheric phenomena and its impact on communications and space weather.

2.3.1 Observational Details

Four experimental campaigns were conducted at HAARP, during the period from October 2009 to November 2010. In all campaigns the HF heater operated at the maximum 3.6 megawatt (MW) power, O-mode polarization, and the beam was directed towards the HAARP MZ. The HF heating frequency was chosen to either match the ionospheric critical frequency (f_0F_2) or coincide with the second electron gyro-harmonic (2.8 MHz). The former frequency was used for daytime ionospheric heating, while the latter frequency was used during nighttime. The modified ionosphere was sensed by instruments aboard DEMETER and by the DMSP satellites available at the time of the experiment. The satellite observations were complemented by ground based diagnostics provided by the HAARP ionosonde and the Kodiak radar.

The French microsatellite DEMETER, which was in operation until December 2010, follows a circular (670 km above-ground) sun-synchronous polar orbit. DEMETER flies over HAARP two times a day, during nighttime at 6 - 7 UT (22 - 23 LT), and during daytime at 20 - 21 UT (12 - 13 LT). The distance between the satellite and the HAARP MZ varies from about 10 km - 500 km due to spin precession. Since we are interested in close overflies of less than 100 km from the HAARP MZ, the number of useful orbits is limited to approximately 1 - 2 per week. Throughout the study we have used the plasma analyzer instrument (Instrument Analyseur de Plasma, IAP) [7] and Langmuir probe instrument (Instrument Sonde de Langmuir, ISL) [27] on-board DEMETER as topside diagnostic tools. At the time of our experiments, the DEMETER instruments operated in "burst" mode,

allowing a sampling rate of 0.43 Hz for the IAP and 1 Hz for the ISL.

The DMSP satellites fly in circular (840 km above-ground) sun-synchronous polar orbits. Although less sensitive than DEMETER, the constellation of DMSP satellites provides HAARP overfly coverage of a few times a day. Restriction of overflies to within 100 km of the HAARP MZ limited the number of useful orbits to about 4 - 5 per week. We used a suite of onboard sensors that measures the ion densities and drift motions of ionospheric ions. The horizontal (V_H) and vertical (V_V) components of the plasma drift velocity have a measurement range of $\pm$ 3 km/s with one-bit resolution of ΔV = 12 m/s, provided $n_i \geq 5000$ cm^{-3}. It takes 4 seconds to sample the ion composition, while the plasma drift and density are sampled at rates of 6 and 24 Hz, respectively.

Approximately 60 DEMETER and DMSP overflies were used in our analysis, which occurred during different daytime conditions and seasons, as well as under quiet and perturbed ionospheric conditions. Table 2.1 shows a summary of relevant information for each heating experiment that detected artificial ionospheric ducts. Columns 1 - 7 of Table 2.1 are: the date of the experiment and name of the satellite used; HF heating time; applied heating frequency (f_H) and reflection height h (the absence of h in the table means that $h = h_mF_2$); ionospheric critical frequency (f_0F_2) and critical height (h_mF_2); the closest approach to the HAARP MZ (ΔR) and E-W half-power beam width at the heating altitude (L_{EW}); the peak change in the field-aligned ion velocity observed by the satellite (ΔV_i); and the relative deviation of the ion density in the duct observed by the satellite ($\Delta n_i/n_i^0$). L_{EW} is assumed to be equal to the size of the heated region. Note that some experiments, marked

Satellite / Date	Heating time (UT)	f_H (MHz) / h (km)	f_0F_2 (MHz) / h_mF_2 (km)	ΔR (km) / L_{EW} (km)	ΔV_i (m/s)	$\Delta n_i/{n_i}^0$ (%)	Kodiak (reflections)	AL (nT)
DEMETER 10/16/09	20:15 - 20:45	5.1 (CW)	5.0-5.1 / 210	69/32	N.A.	11	Strong	10
DEMETER 10/19/09	20:00 - 20:30	5.4	5.4-5.8 / 200	27/28	N.A.	5	Moderate	20
DEMETER 10/21/09	06:15 - 06.30	2.8 (CW)	2.0 / 210	27/58	N.A.	21	N.A.	3
DMSP F15 02/04/10	02:10 - 02:29	4.0	4.0 / 240	45/46	300	30	N.A.	5
DMSP F16 02/09/10	17:40 - 17:59	2.8 / 200	3.4 / 240	65/55	70	8	N.A.	3
DMSP F16 02/10/10	03:30 - 03:50	2.85 / 220	3.6 / 250	10/60	250	40	N.A.	5
DEMETER 02/10/10	20:15 - 20:34	4.25	5.2-5.5 / 210	40/39	N.A.	3	N.A.	100
DMSP F15 02/11/10	02:00 - 02:19	4.25 (CW)	4.9 / 230	25/41	80	8	N.A.	15
DEMETER 06/17/10	20:15 - 20:35	2.85 / 150	4.0-4.5 / 180	141/40	N.A.	5	Weak	250
DMSP F16 06/19/10	03:20 - 03:40	2.85 / 150	4.2-4.5 / 220	54/40	20	3	N.A.	40
DMSP F18 06/24/10	18:45 - 19:00	5.2	4.1-5.2 / 180	76/27	50	4	Strong	50
DMSP F16 06/26/10	17.25 - 17:45	4.3	4.0-4.4 / 220	56/40	30	5	Strong	250
DEMETER 10/28/10	19:55 - 20:15	5.5 (CW)	5.2-5.9 / 220	28/30	N.A.	8	N.A.	100
DMSP F15 10/31/10	01:45 - 01:50	5.6 (CW)	5.3 / 230	66/31	70	5	Strong	2
DMSP F16 11/01/10	17:25 - 17:45	4.1	3.7-4.1 / 230	87/43	55	6	Strong	10
DEMETER 11/04/10	20:10 - 20:30	5.4 (CW)	5.1 / 220	151/31	N.A.	5	N.A.	3
DMSP F15 11/07/10	01:32 - 01:44	5.05 (CW)	4.7-5.7 / 240	30/38	100	15	N.A.	4
DEMETER 11/07/10	20:05 - 20:25	6.5 (CW)	6.5-7.0 / 220	56/26	N.A.	13	N.A.	200
DEMETER 11/09/10	06:05 - 06:25	2.8 (CW)	1.8 / 310	46/85	N.A.	22	N.A.	120
DMSP F16 11/10/10	03:20 - 03:40	2.85 (CW)	3.0 / 230	96/62	270	75	N.A.	10

Table 2.1: A summary of relevant information for each heating experiment.

with a "CW" in column 3, indicate Continuous Wave (CW) heating, while for the remainder square modulation of the HF wave was used (the modulation frequency ranged from 0.1 - 0.7 Hz); the average power of the modulated wave is half that of the CW. In column 6 "N.A" indicates that the velocity measurement was either unavailable or inconclusive. Table 2.1 also lists the results of the ground based observations taken during satellite flyovers. Namely, the eighth column reveals if reflections of the HF signals were detected by the Kodiak radar. Here the "strong" and "moderate" reflection indicate respectively 4 - 8 and 2 - 4 dB increase in the reflected signal intensity relative to that reflected from the ambient ionosphere. All the nighttime experiments in this column are marked by an "N.A.", since the Kodiak radar cannot probe a weak nighttime ionosphere; daytime experiments marked by an "N.A." indicate that the Kodiak radar was not available at that time. The ninth

and final column shows the AL geomagnetic auroral electrojet index. Discussing the observational details of every experiment in Table 2.1 would be quite extensive, thus we will focus on only a few representative cases.

Figure 2.6 shows observations made by the DEMETER IAP along its orbit during two different experiments. Namely, Figure 2.6a shows measurements of the O^+ ion density made during nighttime on 21 October, 2009. The closest approach of

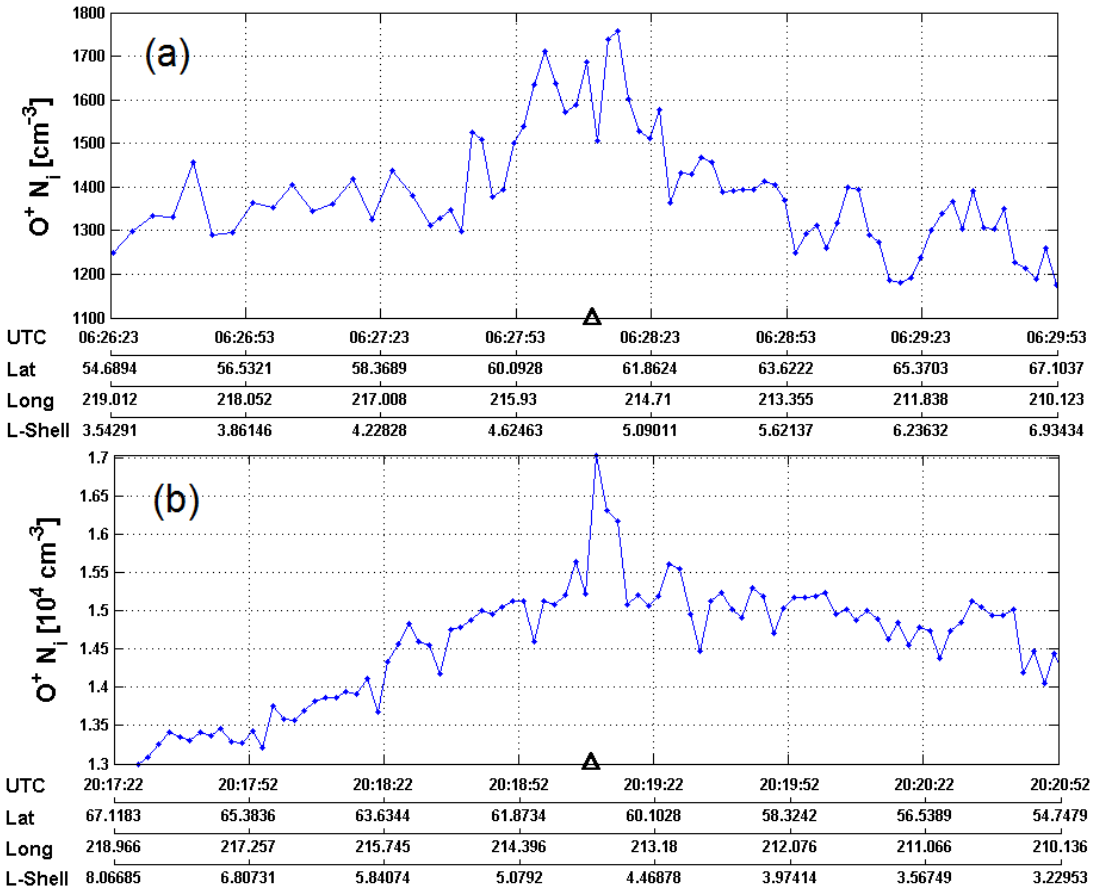

Figure 2.6: DEMETER observations of the O^+ ion density made on 21 October 2009 (a) and 7 November 2010 (b) along its orbit. The multiple scales on the x-axis show the time of observations in UT, satellite (geographic) latitude, longitude, and L-shell; triangle shows the closest approach of DEMETER to the HAARP MZ.

DEMETER to the HAARP MZ occurred at 06:28:08 UT/ 22:28:08 LT (marked by the triangle) and was at a distance of about 27 km. A distinct peak ($\Delta n_{O^+}/n^0_{O^+} \approx$ 21%) was detected when overflying the HAARP MZ, presumably caused by HF

heating-induced plasma transport along the magnetic field. If we define the width of the n_{O^+} peak to be characterized by the full width at half maximum, we find that it is approximately 26 seconds, which corresponds to about 190 km when the DEMETER orbital speed of 7.5 km/s is considered. This is twice the size of the HF heated spot located near the F_2 peak at 230 km, where the size of the HF-heated spot was estimated by taking into account that the half-power beam width at $f_H = 2.8$ MHz is 20.2° in the North-South plane [57]. It should be noted that in the same experiment intense stimulated electromagnetic emission was detected by DEMETER in the HF range [33]. This is an indication of strong ionospheric turbulence due to anomalous absorption near the F_2 peak. Figure 2.6b shows results of a daytime HAARP/DEMETER experiment made on 7 November 2010. A distinct narrow peak of the O^+ ion density ($\Delta n_{O^+}/n^0_{O^+} \approx 13\%$) can be seen when overflying the HAARP MZ. Notice that the peak width is only about 10 seconds, which corresponds to 75 km, i.e. almost 3 times smaller than that during the nighttime. We will discuss this effect later on in the chapter.

Ionospheric modification was also detected by the ISL instrument during this experiment; Figure 2.7 shows the electron temperature (Figure 2.7a) and total electron density (Figure 2.7b). Despite the Langmuir probe often being too noisy to be useful at the auroral latitude of HAARP, in this specific case a fairly distinct peak in electron temperature and density can be seen. As before, the closest flyby of DEMETER to the HAARP MZ is marked by the triangle.

We next discuss the HAARP/F16 DMSP experiments conducted on 9 February 2010 (Figures 2.8a and 2.8b) and 10 November 2010 (Figures 2.8c and 2.8d).

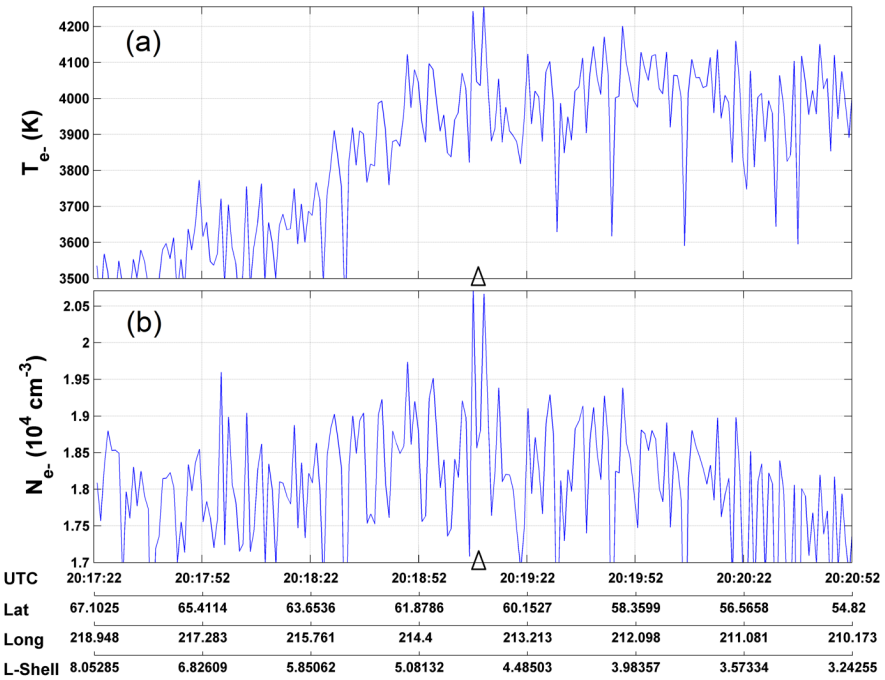

Figure 2.7: DEMETER observations made on 7 November 2010 showing the electron temperature (a) and electron density (b).

Figure 2.8a shows the total ion density in cm^{-3} (solid line), and the O^+ ion density in cm^{-3} ($\times$ markers), while Figure 2.8b shows the upward field-aligned ion velocity, V_{up} (m/s); $V_{up} = V_{vert}/\cos(\alpha)$, where V_{vert} is the vertical ion velocity, and α is the angle between the vertical and the HAARP MZ. Similarly, Figure 2.8c shows the total and O^+ ion densities, while Figure 2.8d shows the upward field-aligned ion velocity. The x-axis gives the time (in seconds) measured relative to the time of the closest approach to the HAARP MZ, T_{mz}. One can clearly see distinct ion outflows of the width $\leq$ 160 km ($\leq$ 20 s) about T_{mz}, which is of the order of the HF-heated spot, and having field-aligned velocities of about 70 m/s on 9 February 2010 and 250 m/s on 10 November 2010; the corresponding relative ion density perturbations are about 8% and 75%, respectively. The large relative increase in density in the latter case can be attributed to the low ambient density of 2000 cm^{-3} and high ion outflow

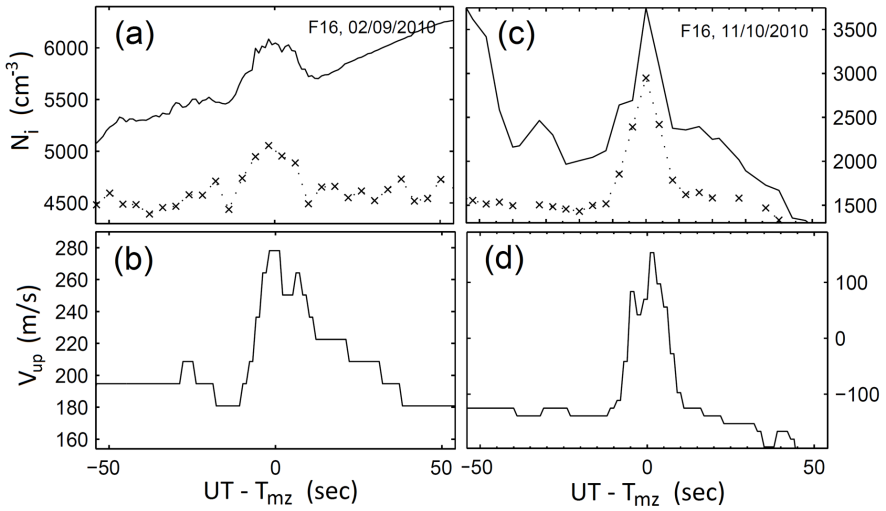

Figure 2.8: Observations made by the F16 DMSP satellite on 9 February 2010 (first column) and 10 November 2010 (second column). The top figures (a and c) show the total ion density in cm^{-3} (solid trace), and the O^+ ion density in cm^{-3} ($\times$ markers). The bottom row shows the upward field-aligned ion velocity in m/s (b and d).

velocity of 250 m/s compared to the 5500 cm^{-3} and 70 m/s of the former case. It should be noted that the spatial profile of the ion outflow is similar to that of the local ion density; namely, they are both bell shaped. This feature can be seen in the previously mentioned DMSP observations as well as in Figure 2.9a, which shows the ion density (top panel) and velocity (bottom panel) perturbations observed by DMSP F15 on 31 October 2010. Figure 2.9b shows the corresponding Kodiak radar observations, which indicate strong reflection of the radar signals during the HF heating at around 01:45 UT. This is indicative of plasma turbulence excitation in the F region plasma, due to anomalous absorption of the HF wave power and creation of plasma density striations that reflect transmitted radar signals.

2.3.2 Model Validation

Now with the experimental results and model description presented, we move on to discussion and modeling results. Note that listed in Table 2.1 is the E-W width

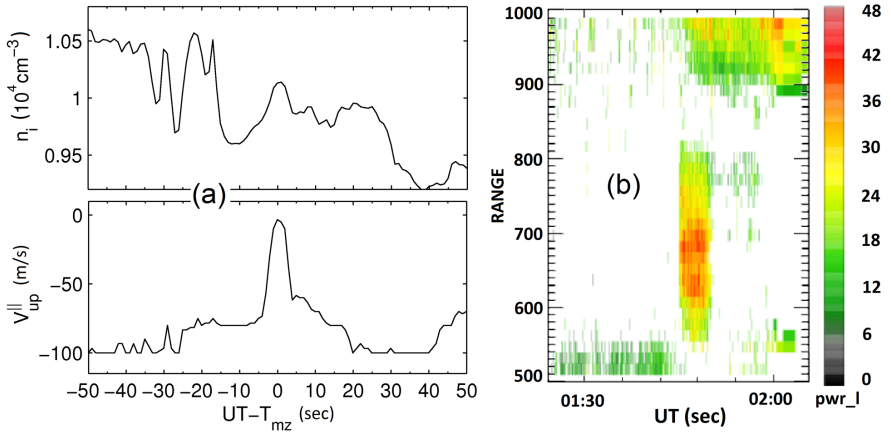

Figure 2.9: Ion density/velocity (a, top and bottom panels, respectively) measured by DMSP and the corresponding Kodiak radar observations (b) made on 31 October 2010 during the HAARP experiment. For the former, time is measured in UT relative to T_{mz} - the time of closest approach to the HAARP MZ.

of the half power beam (L_{EW}) along with the distance of the satellite's orbit from the center of the beam (ΔR). For close flyovers when $\Delta R/L_{EW} < 1$ the probability of duct detection strongly increases along with the measured duct amplitude $\Delta n_i/n_i^0$. During our HAARP experiments, essentially every time a satellite came close to the HAARP MZ while at the same time there being a perceptible ionospheric F_2 peak density, ionospheric ducts were detected. For more distant flyovers the current model should be modified, namely the angular distribution of the HAARP beam should be properly described outside of the half-power region as well.

We now compare model results with the observations described in Section 2.3.1. Figure 2.10 shows comparison of the model results with the DEMETER nighttime and daytime observations of the relative O^+ density n_i/n_i^0 (from here on out we denote n_{O^+} by n_i since $n_{O^+} \approx n_i$ at the relevant altitudes). The values of n_i/n_i^0 measured by DEMETER along its orbit during the nighttime on 21 October 2010 (Figure 2.10a) and during the daytime on 7 November 2010 (Figure 2.10b) are shown by the connected points; for both cases the "ambient" density corresponding

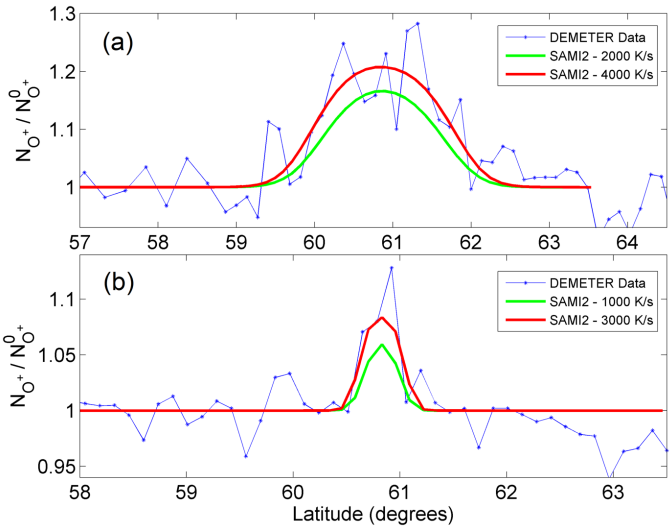

Figure 2.10: DEMETER measurements of the relative perturbations of O^+ density made during nighttime on 21 October 2009 (connected points) along with SAMI2 model results (solid trace) (a); similarly on 7 November 2010, but during daytime (b).

to the DEMETER data was taken to be a constant along the orbit, and equal to the average of the density values to the left and right of the ducts. The curves show model results computed for different T_e pumping rates. Note that due to the difference in the peak plasma frequency during nighttime and daytime, we used different heating frequencies in each case (see Table 2.1). Since the HAARP beam width changes with the transmitted frequency, we adjusted the horizontal size of the heated region accordingly. Namely, at the F_2 peak height (in this case about 220 km) the horizontal sizes for the above nighttime and daytime experiments were estimated to be about 80 km and 36 km, respectively. The figure indicates good agreement between the SAMI2 model predictions and the DEMETER observations. Note that the observed fine structure of the artificial ducts could be caused by large scale irregularities induced by the HF heating which in turn form irregularities inside the ducts when moving upward along the field line.

Figure 2.11 shows the DMSP measurements and results of the SAMI2 model

for the HAARP/DMSP F16 experiments conducted on 9 February 2010. In Figure

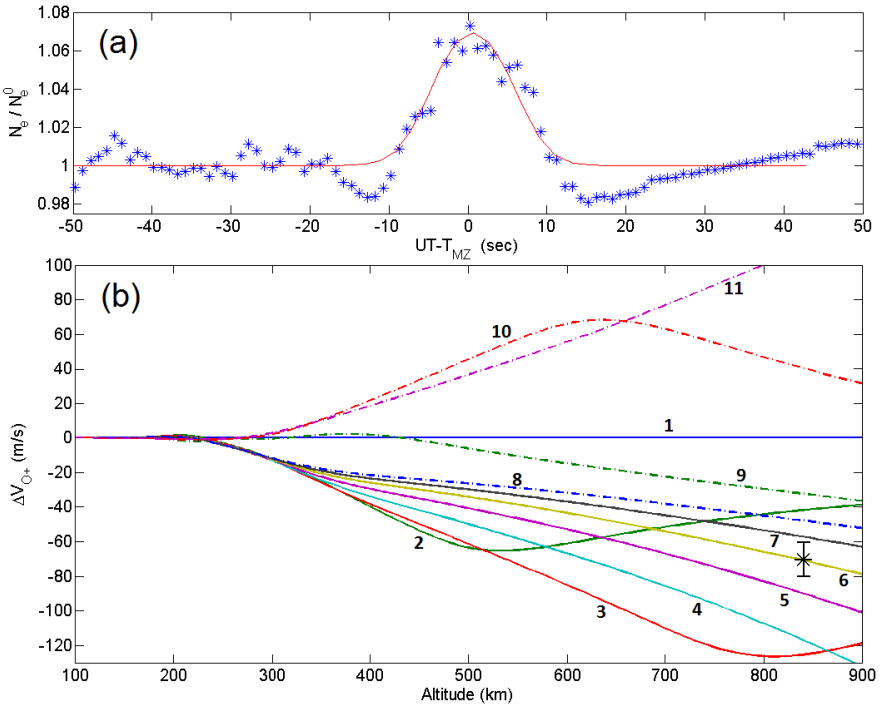

Figure 2.11: SAMI2 modeling for 9 February 2010 HAARP/DMSP F16 experiments. The top panel (a) shows modeled (solid line) and observed (data points) relative electron density. Also shown is the upward ion outflow velocity (b) with several chronologically-numbered traces; the data point with the error bar shows the observation made by F16.

2.11a, the modeled relative ion density is shown by the solid line, while the observations are marked by points having a relative density perturbation $\Delta n_i/n_i^0$ of about 8%. Figure 2.11b shows the SAMI2 computation of the upward ion outflow velocity, where the ambient velocity values have been subtracted away; the time between the chronologically-numbered traces is about 2.5 minutes, and trace 9 represents the first curve after heating has ended. The data point with the error bar toward the right shows the observation of the topside ion velocity (relative to its ambient value) made by the DMSP F16 satellite (column 6 of Table 2.1). It can be seen that the data point is quite close to the black solid trace, corresponding to the velocity profile during the closest of DMSP to the MZ at 17:54:30 UT. The modeling for the

conditions of this experiment was achieved with a heating rate of 700 K/s.

We should mention that all SAMI2 comparisons were done assuming that the satellite trajectories lie in the HAARP magnetic plane. In fact, the magnetic latitude of the satellites varied by 3 - 5° during a given flyby, while the magnetic longitude varied by 1.5 - 4.5°. The deviations from the magnetic plane are often small, but not always. We expect the satellite observations made along their current orbits to be qualitatively similar to observations that would be made if their orbits were in the magnetic plane, especially since SAMI2 comparisons were done with reasonably close flybys ($\Delta R < 60$ km). Properly describing the effects of the heated region geometry and the satellite trajectory would require a 3D model. $\mathbf{E} \times \mathbf{B}$ drifts have also been neglected throughout our modeling. Future work will involve shifting from using SAMI2 to using the 3D model SAMI3, where the above considerations will be taken into account as much as possible.

2.4 Effects of Ducts on Wave Propagation

In addition to large scale plasma transport induced by ionospheric heating, two additional effects were observed during our experiments. The first of these occurred on 16 October 2009, when DEMETER detected a daytime duct of 11% enhancement (see Table 2.1) and simultaneously strong VLF signals near 8 kHz. The VLF was observed only during DEMETER crossing the HAARP MZ, suggesting that they were a result of the ionospheric heating. The generation mechanism for the VLF observed in this experiment will be discussed in great detail in Chapter 4, and thus does not concern us here. Rather we will discuss how the whistler waves are

affected by the presence of the artificial duct. It will be shown that the whistler waves are "ducted" due to the change of index of refraction they experience during their propagation along the geomagnetic field. Simulation results will be presented, based on the work of [59].

The second effect is similar to the first; we have observed what we believe to be the focusing and trapping of the HF pump wave by density depletions in the ionosphere. Focusing of HF requires a drop in plasma density relative to the surroundings (and thus an increase in the refractive index), which is observed during two experiments by DEMETER. Simultaneously, DEMETER detected the HAARP HF signal along with a multiple frequency band structure that spans the entire frequency range. Such a phenomena is an artifact due to DEMETER being exposed to a very intense HF beam, causing a saturation of its instrument. It is shown with simple theoretical considerations that the intensity of the beam necessary to cause such a saturation of the instrument can only occur if the HF waves have been focused by the presence of density cavities (based on the work of [33]).

2.4.1 VLF Ducted Propagation

On 16 October 2009, an ionospheric heating experiment was conducted with HAARP using a 5.1 MHz O-mode wave. DEMETER crossed within 70 km of the HAARP MZ at 20:33:02 UT. Measurements from the HAARP ionosonde taken 2 minutes later (not shown) reveal that the heating frequency was marginally below the peak ionospheric plasma frequency, $f_0F_2 = 5.15$ MHz, which occurred at an altitude of $h_mF_2 = 225$ km. The altitude of peak heating (where the heater frequency

matches the local plasma frequency) was just slightly below $h_m F_2$, occurring at about 215 km.

As shown in Figure 2.12, onboard measurements indicate the presence of multiple regions of enhanced electron density as well as increased VLF wave power between 7 and 10 kHz. Notably, there is an apparent correlation between increased electron density and enhanced VLF power. There are three electron density enhancements in Figure 2.12b; we refer to these as "Duct 1" ($t = 0$ s), "Duct 2" ($t \approx$ -3.5 s), and "Duct 3" ($t \approx$ -8 s). Comparing Figures 2.12a-b, we see that Ducts 1 and 3 coincide with the two most prominent spectral enhancements, and Duct 2 corresponds to a weaker enhancement that maximizes around 8 kHz. Between Ducts 1 and 2 is a somewhat stronger spectral enhancement that does not have a corresponding region of enhanced density.

To study how the presence of the density structures affect the whistlers, Reference [59] used an Electron MagnetoHydroDynamics (EMHD) model to compute the whistler wave propagation. The primary equations of EMHD are Faraday's Law, Ampère's Law, and the (linearized) momentum equation for an electron fluid:

$$\frac{\partial \mathbf{B}}{\partial t} = -\boldsymbol{\nabla} \times \mathbf{E} \tag{2.17}$$

$$\boldsymbol{\nabla} \times \mathbf{B} = -\mu_0 n e \mathbf{u} \tag{2.18}$$

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{e}{m_e}(\mathbf{E} + \mathbf{u} \times \mathbf{B}_0) - \nu \mathbf{u}, \tag{2.19}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields, $\mathbf{u}$ is the electron fluid velocity, n is the electron (number) density, and ν is the total electron collision frequency. Also, note that Equation (2.18) assumes that the current is due entirely to electron

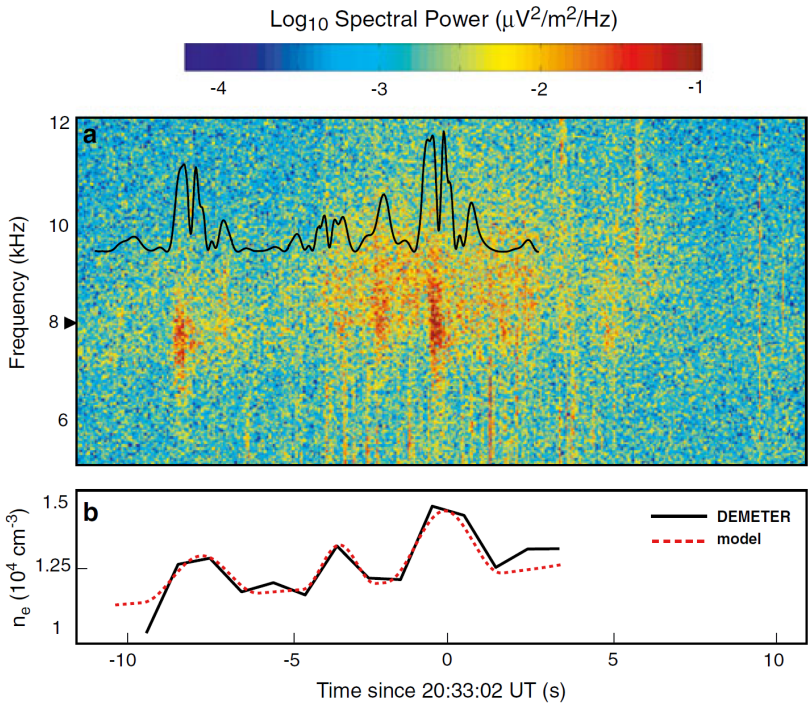

Figure 2.12: (a) Time-frequency spectrogram of electric field observed by DEMETER. The black line shows a schematic profile of electric field intensity from the EMHD model. (b) Plasma density as measured by the DEMETER and the density profile used in the model.

motion, $\mathbf{J} = -ne\mathbf{u}$, since ions are expected to be immobile on the whistler wave timescales. Equations (2.18) and (2.19) can be combined to obtain Ohm's law for the electric field,

$$\lambda_e^2 \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E} + \mathbf{E} = -\frac{m_e \nu}{e}\mathbf{u} - \mathbf{u} \times \mathbf{B}_0, \tag{2.20}$$

where $\lambda_e \equiv c/\omega_p$ is the electron inertial length. The model is solved in magnetic dipole coordinates. The electric field is obtained from Ohm's law (2.20) using an iterative relaxation method, and the magnetic field in (2.17) is time-stepped using a fourth-order predictor-corrector approach. The momentum equation (2.19), is not explicitly solved in this model, and the fluid velocity is instead determined directly from the curl of the magnetic field via Ampère's Law (2.18).

Using a density profile that matches the DEMETER duct observations (Fig-

ure 2.12b) the model can be solved to yield the wave intensity profile. Figure 2.13a shows the spatial distribution of electric field amplitudes in our simulation of this event. Initially, the electric fields inside each of the ducts are transversely localized

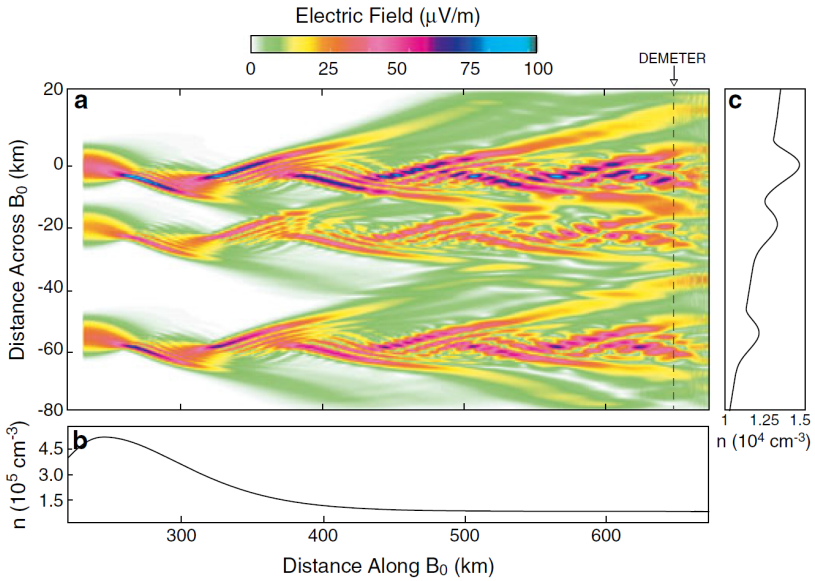

Figure 2.13: Simulation of whistler propagation in model density ducts. (a) Distribution of electric field amplitude in the simulation domain. The trajectory of the DEMETER satellite is indicated by a dashed black line. (b) Background field-aligned density profile at the center of the topmost duct. (c) Density profile at DEMETER altitude.

near their source. However, as the waves propagate along the field line they leak from their ducts and eventually electric field power becomes distributed throughout the entirety of the domain. The intensity of the electric field measured by DEMETER is shown schematically in Figure 2.12a. As can be seen in the figure, there is very good agreement between the spatial distribution of electric fields observed by DEMETER and the simulation. Notably, the spectral peak that occurs between Ducts 1 and 2 is present, and it has an amplitude greater than the wave in Duct 2. The constructive interference of leaked wave power from Ducts 1-2 is apparently responsible for producing the observed enhancement. Further analysis shows that the leakage is caused by the narrow and asymmetric nature of the ducts [59].

2.4.2 HF Focusing

In this section we present the first direct evidence of HF focusing induced by ionospheric ducts, along with a simple theoretical model. The experiments were conducted by injecting HF radio waves into the F region ionosphere using the HAARP heater, and detected by instruments on board the DEMETER satellite. DEMETER observed a multiple frequency band structure characteristic of the strong HF signal exceeding the saturation level of the detector. Analysis of the O^+ density measured by DEMETER along its orbit shows that the strong radio signal coincides with the presence of a "negative" duct in the ionosphere (ionospheric trough); "negative" refers to the presence of a plasma density depletion with the peak depletion located near the center of the duct. Such ducts constitute a change in the index of refraction leading to the focusing of HF waves in a manner equivalent to a "thick" plasma lens. The way in which a density depletion can affect HF propagation is illustrated by Figure 2.14, which shows a schematic of HF focusing by a negative ionospheric duct. The change in the index of refraction acts as a lens and bends the HF beams

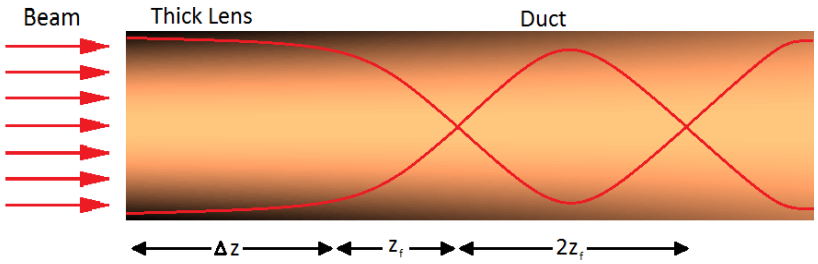

Figure 2.14: Schematic of HF wave focusing by an ionospheric density depletion. Areas of darker (lighter) colors represent higher (lower) plasma density. The ray path of the radio waves is shown and demonstrates the formation of multiple nodes.

toward the center. Outside the lens, the focused beam propagates inside the duct and reflects from the duct walls, and as a result produces focal nodes. In what fol-

lows we present the first experimental evidence of ionospheric HF focusing by such a plasma lens, along with a simple analytic model of this process.

2.4.2.1 Observations

A set of experiments were performed in which the HAARP heater injected 3.6 MW (84.1 dBW ERP) of O-mode radio wave power along the MZ at a frequency of 2.8 MHz; MZ heating and 2.8 MHz were chosen to closely match the f_0F_2, because previous experiments indicated that they maximize the probability for duct formation. The DEMETER satellite was used as a diagnostic during its close flyby to the HAARP MZ. The previously mentioned IAP instrument (see Section 2.3.1) was one of two DEMETER diagnostic instruments used in the experiments discussed below; it was mainly used for density measurements of O^+ ions - the most abundant ions in the F region. The second instrument is the ICE (Instrument Champ Electrique), which measures HF signals in a frequency range from a few kHz to 3.3 MHz [6]. As usual, the local ionospheric conditions were monitored by the HAARP ionosonde.

The HF spectrogram observed by the ICE on 02/12/10 at 6:29-6:33 UT is shown at the bottom of Figure 2.15. In this experiment the closest distance of DEMETER from the MZ was 38 km. The observed spectral "line" at 2.8 MHz[12] that extends between 40.3° and 65° latitude - over 2800 km - is generated by radio emission stimulated by the interaction of the injected HF with the F region plasma, rather than by the direct "free space" HAARP beam.

We now focus our attention on the strong multiple-frequency band structure

[12]It should be mentioned that the maximum detection frequency of the ICE is 3.3 MHz (the sampling frequency is 6.6 MHz), thus only nighttime experiments tend to show the HAARP beam since frequencies above 4 MHz are required to match the daytime f_0F_2.

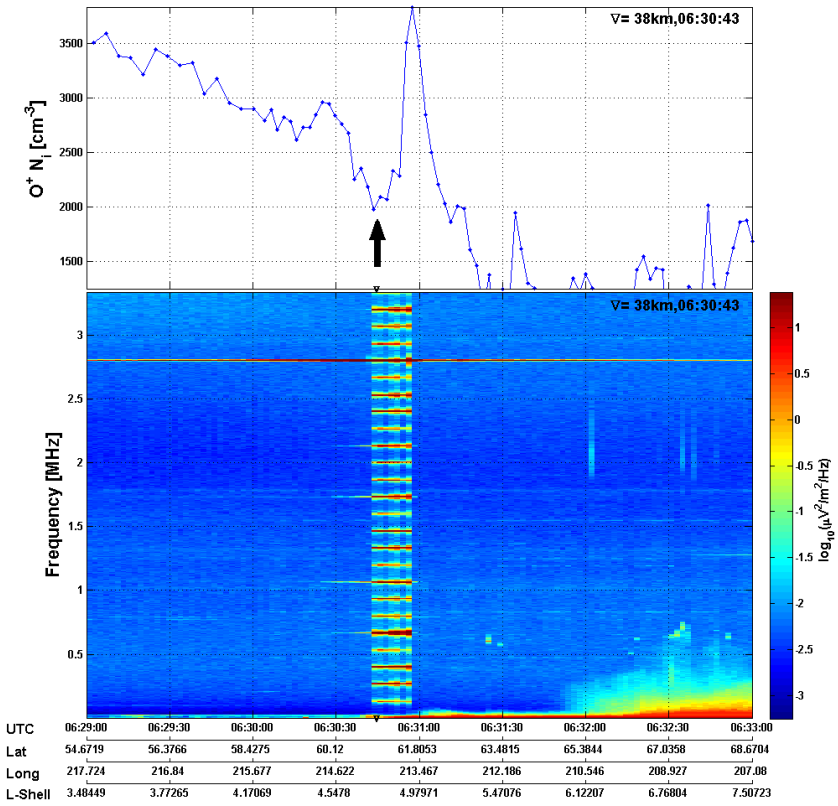

Figure 2.15: DEMETER observations of the O^+ density (top) made on 02/12/10, and simultaneously the HF electric field power spectrum (bottom). The arrow points to the minimum of the relevant negative duct.

observed between 6:30:40 and 6:30:50 UT. Such a structure is indicative of the ICE receiving a strong HF signal exceeding its saturation level, which at 2.8 MHz is approximately 10 mV/m. Analysis of the O^+ density measured by DEMETER along its orbit (see the top of Figure 2.15) shows that the strong signal detected by ICE almost coincides with the presence of a negative duct in the ionosphere. Plasma quasineutrality in the ionosphere requires that the electron density exhibit a similar profile. It is well known that a depletion in electron density leads to a positive perturbation in the index of refraction. Thus incident HF waves propagating into a plasma channel with a field-aligned density depletion will be focused towards the center (Figure 2.14).

Figure 2.16 shows another example of HF focusing observed on 10/21/09 (same experiment from Figure 2.6a). The ionosphere was quiet with the F_2 peak at 220

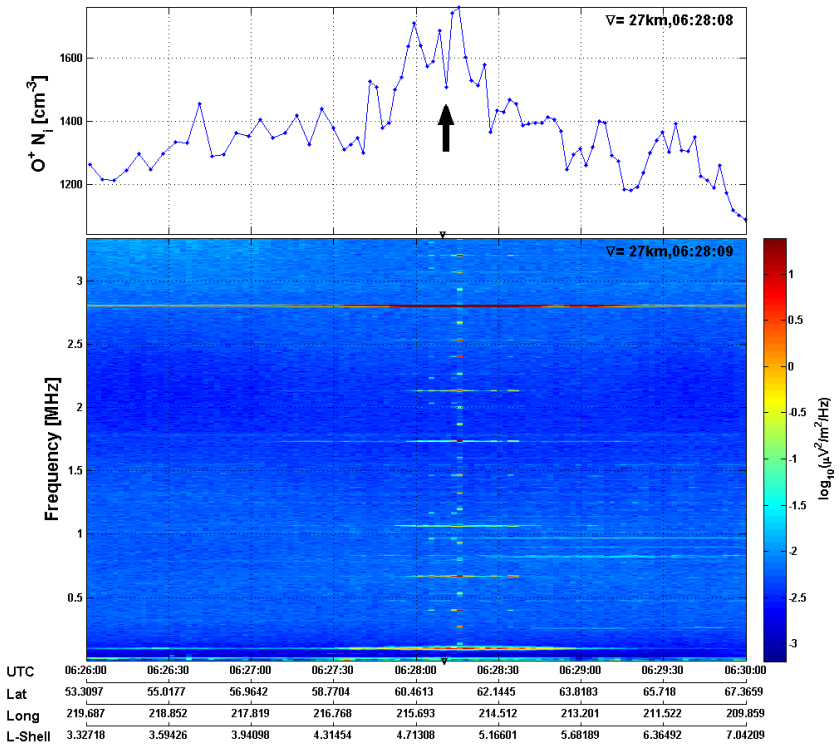

Figure 2.16: Same as Figure 2.15 but during a difference experiment.

km and $f_0F_2 = 2.05$ MHz. The ICE HF spectrogram in Figure 2.16 (bottom) shows a band at about 06:28:15 in the entire frequency range of the instrument. Figure 2.16 (top) shows the measured O^+ density. The presence of a negative duct within the artificially created positive duct[13] can be seen (arrow), albeit not as distinct as in Figure 2.15.

2.4.2.2 Theoretical Model

A simple theoretical model can be formulated by considering a plane wave with (angular) frequency ω propagating along the field aligned density depletion shown

[13]This case is rather different in that regard. The density depletion in Figure 2.15 is natural, while the one in Figure 2.16 is a conventional HAARP induced artificial duct, but with a relative depletion *inside*.

in Figure 2.14. The index of refraction at F region altitudes can be approximated as Equation (1.15): $\eta = \sqrt{1 - \omega_p^2/\omega^2}$. In the presence of the density duct the index of refraction is perturbed due to the change in the electron density. Assuming a density perturbation Δn with $\Delta n/n << 1$ the perturbed index of refraction $\Delta\eta$ can be expressed as

$$\Delta\eta = \frac{\partial\,\eta}{\partial\,n}\Delta n = -\frac{1}{2}\frac{\omega_p^2/\omega^2}{\sqrt{1-\omega_p^2/\omega^2}}\frac{\Delta n(z,\rho)}{n(z)}.$$

Here $n(z)$ is the ambient electron density. Following Reference [20], the electric field at distance z_1 from the lower boundary of a duct and radius ρ is given by

$$E = Ae^{i\psi_0}\frac{\omega}{ic}\int\frac{1}{r}e^{i\Delta\phi(z_1,\rho)+i\frac{\omega}{c}r}\rho\;d\rho. \tag{2.21}$$

Here A is the amplitude of the wave at the lower boundary of the duct, ψ_0 is an arbitrary phase, and $r = \sqrt{z_1^2+\rho^2}$. The perturbed phase of the wave at a distance z_1 along the duct is given by

$$\Delta\phi = \frac{\omega}{c}\int_0^{z_1}\Delta\eta\;dz.$$

In the absence of the phase perturbation $\Delta\phi = 0$ and $|E| = A$. Assuming the duct is smooth and cylindrically symmetric, we expand $\Delta\phi$ in powers of ρ^2:

$$\Delta\phi = \Delta\phi_0 + \rho^2\Delta\phi_1 + \rho^4\Delta\phi_2 + \cdots$$

and taking into account that $\rho/z_1 << 1$, i.e. that the duct has a limited transverse size, we find that

$$r = z_1 \left(1 + \frac{\rho^2}{2z_1^2} + \cdots\right).$$

Focusing occurs when the phase in the exponent of Equation (2.21) approaches zero [20]. Thus, to second order in ρ/z_1, the focal distance z_f satisfies the equation

$$\Delta\phi_1 + \frac{\omega}{2cz_f} = \frac{\omega}{2c}\int_0^{z_f} \frac{\omega_p^2/\omega^2}{\sqrt{1-\omega_p^2/\omega^2}} \frac{1}{n}\frac{\partial\,\Delta n}{\partial(\rho^2)} dz - \frac{\omega}{2cz_f} = 0. \tag{2.22}$$

Solving Equation (2.22) for z_f we find that

$$z_f = \left\{ \int_0^{z_f} \frac{\omega_p^2/\omega^2}{\sqrt{1-\omega_p^2/\omega^2}} \frac{1}{n}\frac{\partial\,\Delta n}{\partial(\rho^2)} dz \right\}^{-1}. \tag{2.23}$$

From this point forward we will take the above expression and evaluate it using certain assumptions that are tailored to each of the two experiments described above. Focusing our attention first to the experiment performed on 2/12/2010, we assume for simplicity that the radial size of the duct, ρ_0, has a weak dependence on altitude (or none at all). This will allow treating ρ_0 as a constant and pulling it out of the integral. In addition we note that the product of ω_p^2/ω^2 and $1/n$ is a constant, and assume that the electron density above the F_2 peak can be modeled as a decaying exponential: $\omega_p^2/\omega_0^2 = n/n_0 = \exp(-z/H)$, where ω_0 and n_0 are the plasma frequency and density at the F_2 peak, respectively, and H is to be chosen in such a way that $n(z_{sat})$ gives the ambient electron density measured by

the DEMETER satellite. After these simplifications Equation (2.23) becomes

$$\frac{1}{z_f} = \frac{\omega_0^2}{\omega^2}\frac{1}{\rho_0^2}\int_0^{z_f}\frac{\Delta n(z)}{n_0}\frac{dz}{\sqrt{1-(\omega_0^2/\omega^2)e^{-z/H}}}. \tag{2.24}$$

The functional form of Δn remains to be specified, and to do so we use EISCAT observations of a quiet-time ionospheric trough. The observations come from Reference [58], in which the electron density of an ionospheric trough was measured at EISCAT at different altitudes, and it was shown that the density depletion (Δn) decreases with altitude. Assuming that the decrease is exponential, the functional form of Δn becomes $\Delta n = \Delta n_0 \exp(-\alpha(z/H))$, where $\Delta n_0 = \Delta n_{sat}\exp(\alpha(z_{sat}/H))$ is the density depletion measured at the F_2 peak, n_{sat} is measured by the DEMETER satellite, and α is the decay constant to be quoted later. Finally, we introduce normalized coordinates $\bar{z}_f = z_f/H$ and $\xi = z/H$ and thus Equation (2.24) becomes

$$\frac{1}{\bar{z}_f} = \frac{\omega_0^2}{\omega^2}\frac{H^2}{\rho_0^2}\int_0^{\bar{z}_f}\frac{e^{-\alpha\xi}d\xi}{\sqrt{1-(\omega_0^2/\omega^2)e^{-\xi}}}. \tag{2.25}$$

This equation can be readily solved numerically as soon as the experiment and duct parameters are specified.

We now shift our attention to the focusing observations made on 10/21/2009. This case differs from the previous case since in this case the duct was of artificial origin, and was caused by the plasma outflow moving along the geomagnetic field line from the HF-heated region located near the F_2 peak of the ionosphere [32]. Such artificially generated ducts have a nearly constant radius and density perturbation (ρ_0 = constant and $\Delta n = \Delta n_0$ = constant). Moreover the term in the denominator

in Equation (2.25) deviates from unity by less than 10% for this specific case, and thus can be safely replaced by unity. After these assumptions the expression for the focal length becomes

$$\frac{1}{\bar{z}_f} = \frac{\omega_0^2}{\omega^2}\frac{H^2}{\rho_0^2}\int_0^{\bar{z}_f} d\xi,$$

which when solved for the focal length gives

$$z_f = \rho_0 \frac{\omega}{\omega_0}\sqrt{\frac{n_0}{\Delta n(z_{sat})}}. \tag{2.26}$$

From Equation (2.21) the amplification of the electric field amplitude caused by focusing is given by

$$\frac{|E|}{A} = \frac{\omega}{c}\frac{\rho_0^2}{z_f}, \tag{2.27}$$

which is valid for either of the experiments.

Note that Equations (2.26) and (2.27) are similar to focusing and magnification by optical lenses. The focal distance is proportional to the lens aperture and inversely proportional to the optical density of the lens material. The lens material (refractive index) and frequency control the magnification coefficient.

We proceed now to use the observations in conjunction with the theoretical model to examine consistency of the expected focal length and magnification with the observations. Referring first to the 02/12/2010 experiment (Figure 2.15), we note that the horizontal size of the duct at the DEMETER altitude is about 170 km, i.e. $\rho_0 = 85 \pm 15$ km (see appendix of Reference [33] for details on uncertainty); the relative depletion of the plasma density inside the negative duct is $\Delta n_{sat} = 1000 \pm 150$ cm^{-3}, while the unperturbed density immediately left of the duct is about

$n_{sat} = 2950 \pm 50$ cm^{-3}. The value of z_{sat} is simply 670 - 300 = 370 km, while H was found to be about 110 km, and the F_2 peak density n_0 can be computed using the value of $\omega_0 = 2\pi \times 2.55$ MHz from the HAARP ionosonde. Based on the density data in [58], the decay constant was estimated to be $\alpha = 0.6 \pm 0.1$ (see appendix of Reference [33]). On the basis of these estimates, corresponding uncertainty ranges, and Equation (2.25) we can estimate the focal length of the focusing duct as about 430 ± 75 km. Thus the duct whose lower boundary is located around the F_2 peak at 300 km provides optimal focusing at about 730 ± 75 km. From Equation (2.27) we find that a wave having frequency $f = 2.8$ MHz is magnified by 940 ± 280 times at the focal point. We can estimate the magnification at the DEMETER altitude by assuming the focused beam has a conical shape, thereby decreasing the magnification by a factor $((z_f - d)/z_f)^2$, where d is the distance between the focal point and the DEMETER orbit. The result is a magnification of about 690 ± 310 at the DEMETER altitude. Our estimates show that during its pass on 02/12/10, DEMETER was located within the focal point uncertainty range of the "lens" formed by the negative duct in the ionosphere. Therefore its antenna received a strongly amplified signal. Furthermore, considering that the power density detected by ICE outside of the duct was 25 $(\mu\text{V/m})^2/\text{Hz}$, and considering that the half bandwidth of the signal is about 12 kHz, we find that the strongest signal outside of the perturbed region was about 0.5 mV/m. Since satellite calibration tests prior to flight gave a 10 mV/m saturation level at 2.8 MHz, the observations of Figure 2.14 require a focusing of at least by 20, which is within the uncertainty range of the magnification and is thus consistent with the observations.

Referring next to the 10/21/09 case, we note that the top of Figure 2.16 shows the clear detection of an artificial duct. Moreover, at 06:28:10 the bottom of Figure 2.16 shows a strong band in the spectrogram which coincides with a local negative duct (marked by an arrow) having a peak density depletion of $\Delta n = 150 \pm 20$ cm^{-3} and radius $\rho_0 = 17 \pm 3$ km [33]. Following Equation (2.26), the focal length of the focusing duct is about 430 km from the F_2 peak, producing 41 ± 16 times magnification at the focal point of this lens (see Equation (2.27) with $f = 2.8$ MHz). We can estimate the magnification at the DEMETER altitude by our previous method of assuming a conical shape for the focused beam. The result is a magnification of about 31 ± 12 at the DEMETER altitude.

2.5 Conclusions

Artificial ducts in the topside ionosphere over HAARP have been detected on a regular basis by the DEMETER and DMSP satellites during four HAARP/BRIOCHE campaigns in 2009 and 2010. Overall, the plasma density enhancements detected between 09:30 and 12:30 LT varied from 3% - 13%, while those during 17:30 - 22:15 LT were typically 15 - 40%. The duct magnitudes and widths provided by the modified SAMI2 model agree fairly well with the satellite observations. The ducts seem to be produced most efficiently for heating frequencies quite close to the critical frequency f_0F_2. Moreover, the ducts' amplitudes seem to be larger for cases where the topside ambient density is lower. This is simple to understand since a lower background density allows density perturbations to stand out more from the natural noise; loosely speaking, adding 1 to 10 is a bigger change than adding 1 to

100. This effect can be seen by comparing typical DEMETER observations with typical DMSP observations; DMSP having an orbit higher by ~ 200 km has a lower local plasma density and thus tends to detect greater values of $\Delta n/n$ during duct experiments. Generation of ionospheric ducts has the potential to facilitate Radiation Belt Remediation (RBR) and significantly increase the lifetime of satellites. Namely, whistler waves can be injected from the ground (or ionosphere) and guided by these ducts into the radiation belts without a significant loss of amplitude, subsequently leading to the precipitation of high energy electrons by resonant pitch angle scattering. Two additional interesting effects having potential applications were observed by DEMETER during the HAARP/BRIOCHE campaigns: ducted whistler wave propagation and HF pump wave focusing. The ducted whistler wave propagation was modeled by [59] and shown to be in good agreement with the amplitude of the electric field observed by DEMETER. Observations were presented of what we believe to be the first direct evidence of HF focusing induced by natural and artificial ionospheric ducts. A simple theoretical model was presented which showed that the presence of a plasma density depletion can lead to focusing at LEO satellite altitudes with enough magnification to cause instrument saturation, thus demonstrating consistency with the observations. Focusing of HF waves could have potential communication and military uses.

Chapter 3

Effects of Ionospheric Heating: Plasma Wave Generation

Air is a familiar example of a neutral gas that can support traveling disturbances of pressure and density in the form of sound waves. It comes as no surprise that a plasma, being an ionized gas, can support waves as well. The major difference, however, is that a plasma contains charged particles, thus giving it a high electrical conductivity. This has far-reaching consequences, which can be appreciated by picturing how the motion of charged particles can affect the plasma. Namely, charged particles respond to a driving electric field, thus generating electric currents and magnetic fields. Moreover, temporal changes of the magnetic field can induce an electric field, and thus separation of positive and negative charges. Together the induced fields can alter the dynamics of the charged particles (and thus the original current) to create new fields, thus repeating the process all over again. This feedback can create a slew of phenomena in the form of numerous types of plasma waves and plasma instabilities, of which a subset of the former will be the focus of this chapter. In fact, plasmas can support a variety of sound waves, electrostatic waves, electromagnetic waves, and combinations thereof. In addition, some wave types can transform into others via a process called *mode conversion*.

Arguably the characteristic behavior of plasma waves is entirely determined by their spatial and temporal scales. For ionospheric conditions, ULF ($\lesssim$ 10 Hz) and ELF ($\sim$ 10 - 1000 Hz) plasma waves have long time scales (i.e wave periods) and are thus associated with the slow ion motion in the plasma. On the other hand, VLF ($\sim$ 1 - 10 kHz) and HF ($\sim$ 1 - 30 MHz) waves have much shorter wave periods, and are associated with electron motion, or a hybrid of electron/ion motion (for VLF). As a review and segue to the latter parts of the chapter, the next section will introduce some of these waves in a mathematical context by working out the plasma dielectric tensor. It is shown that the principle electromagnetic plasma wave modes in the ULF range are the shear Alfvén and magnetosonic waves. The purpose of this chapter is to explore how these two wave modes can be generated by *modulated ionospheric heating* and how they subsequently propagate in the ionosphere and penetrate toward the ground.

3.1 General Overview

A large portion of the types of waves that can exist in plasmas can be worked out by Fourier analyzing the fluid-Maxwell system of equations. Since Fourier analysis is only meaningful for linear systems, the equations must first be linearized about an equilibrium. This is often a valid assumption as wave amplitudes are small in most situations to be considered in this thesis. In addition, it will be assumed here that thermal effects are negligible compared to electromagnetic effects. In other words, the pressure term ($-\boldsymbol{\nabla} p$) in the momentum equation is assumed small[1] com-

[1]This too is often valid, since EM forces tend to dominate thermodynamic ones; the relevant exception is the heated region where the HF pump wave is depositing its energy.

pared to the Lorentz force - the so-called *cold plasma approximation*. The purpose of this chapter is to review the *cold plasma dielectric tensor*, and to introduce wave modes that are relevant to the remainder of the chapter.

3.1.1 Waves in Plasmas

The equations that comprise the fluid-Maxwell system are Faraday's law, the Ampère-Maxwell law, and the continuity and momentum fluid equations for each plasma species (α). The background equilibrium (zeroth order) plasma is assume to be homogeneous, and for simplicity collisions are neglected. Thus the cold plasma equations (before linearization) are given by

$$\frac{\partial n_\alpha}{\partial t} + \boldsymbol{\nabla} \cdot n_\alpha \mathbf{v}_\alpha = 0, \tag{3.1a}$$

$$m_\alpha n_\alpha \left(\frac{\partial}{\partial t} + \mathbf{u}_\alpha \cdot \boldsymbol{\nabla} \right) \mathbf{u}_\alpha = q_\alpha n_\alpha (\mathbf{E} + \mathbf{u}_\alpha \times \mathbf{B}), \tag{3.1b}$$

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \sum_\alpha q_\alpha n_\alpha \mathbf{u}_\alpha + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}, \tag{3.1c}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{3.1d}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0, \tag{3.1e}$$

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{1}{\epsilon_0} \sum_\alpha q_\alpha n_\alpha, \tag{3.1f}$$

where $\sum_\alpha q_\alpha n_\alpha \mathbf{u}_\alpha$ is the current density, $\mathbf{J}$. Equations (3.1e) and (3.1f) work as initial conditions for Equations (3.1c) and (3.1d), since the temporal evolution of the former can be obtained by taking the divergence of the latter. The assumption of homogeneity above means that, to zeroth order, the density $n_{\alpha,0}$ and geomagnetic field $\mathbf{B}_0$ are constant; all other equilibrium quantities are assumed to be zero.

With the above definitions, and denoting perturbed (first order) quantities with the subscript 1, the dynamic variables in (3.1) can be written as

$$n_\alpha = n_{\alpha,0} + n_{\alpha,1}; \quad \mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1; \quad \mathbf{E} = \mathbf{E}_1; \quad \mathbf{u}_\alpha = \mathbf{u}_{\alpha,1}.$$

Note that to zeroth order, Equations (3.1) are automatically satisfied. Substituting these expressions into (3.1) and keeping only terms of first order[2] we obtain the linearized equations for a cold plasma:

$$\frac{\partial n_{\alpha,1}}{\partial t} + n_{\alpha,0}\boldsymbol{\nabla}\cdot\mathbf{u}_{\alpha,1} = 0, \tag{3.2a}$$

$$\frac{\partial \mathbf{u}_{\alpha,1}}{\partial t} = \frac{q_\alpha}{m_\alpha}(\mathbf{E}_1 + \mathbf{u}_{\alpha,1}\times\mathbf{B}_0), \tag{3.2b}$$

$$\boldsymbol{\nabla}\times\mathbf{B}_1 = \mu_0\sum_\alpha q_\alpha n_{\alpha,0}\mathbf{u}_{\alpha,1} + \frac{1}{c^2}\frac{\partial \mathbf{E}_1}{\partial t}, \tag{3.2c}$$

$$\boldsymbol{\nabla}\times\mathbf{E}_1 = -\frac{\partial \mathbf{B}_1}{\partial t}, \tag{3.2d}$$

$$\sum_\alpha q_\alpha n_{\alpha,0} = 0, \tag{3.2e}$$

where the divergence of the magnetic field ($\boldsymbol{\nabla}\cdot\mathbf{B}_1 = 0$) has been omitted for brevity, since it will not be explicitly used in what follows. Incidentally, note that Equation (3.2e), corresponding to a perturbative analysis of Gauss's law (3.1f), yields the quasineutrality condition: $\sum_\alpha q_\alpha n_{\alpha,0} = 0$; this can be the mathematical justification for quasineutrality[3], if you wish. Since the pressure term is neglected in Equation

[2]To be proper the expressions for n, $\mathbf{B}$, $\mathbf{E}$, and $\mathbf{u}$ should have a small parameter, say, ϵ that distinguishes the order of each term: e.g. $n_\alpha = n_{\alpha,0} + \epsilon n_{\alpha,1} + \epsilon^2 n_{\alpha,2}\cdots$. What we are really doing is plugging these kinds of expressions into the equations and collecting like powers of ϵ and truncating the series after first order; this is nothing more than (first order) perturbation theory.

[3]For a singly ionized plasma the charges are $q_e = -e$, $q_i = e$, and we recover the familiar condition $n_{e,0} = n_{i,0}$. Of course, the full densities (n_e, n_i) are the zeroth order quantities plus higher order terms, which is why we often write $n_e \simeq n_i$.

(3.2b), the continuity equation (3.2a) is decoupled from the rest of the equations and will be left out of the system of equations in what follows. The details of finding the wave properties contained in Equations (3.2) are left to Appendix B.1, but the end result is the *cold plasma dispersion relation* (CPDR) expressed in terms of the index of refraction η:

$$\eta^2 = \frac{B \pm F}{2A},$$

where

$$A \equiv S\sin^2\theta + P\cos^2\theta; \quad B \equiv RL\sin^2\theta + PS(1+\cos^2\theta); \quad C \equiv PRL,$$

$$F^2 \equiv B^2 - 4AC = (RL - PS)^2\sin^4\theta + 4P^2D^2\cos^2\theta,$$

and S, D, P, R, and L are defined by Equations (B.7); the background field $\mathbf{B}_0$ is taken to be in the $\hat{\mathbf{z}}$ direction and θ is the angle between the wave vector $\mathbf{k}$ and $\mathbf{B}_0$. An equivalent form of the above dispersion relation [54] is given by

$$\tan^2\theta = \frac{-P(\eta^2 - R)(\eta^2 - L)}{(S\eta^2 - RL)(\eta^2 - P)}. \tag{3.4}$$

The dispersion relations for the special cases of propagation at $\theta = 0$ and $\theta = \pi/2$ are easily obtained from (3.4). For *parallel* ($\theta = 0$) and *perpendicular* propagation ($\theta = \pi/2$) we have

$$P = 0, \quad \eta_\parallel^2 = R, \quad \eta_\parallel^2 = L; \quad \eta_\perp^2 = \frac{RL}{S}, \quad \eta_\perp^2 = P. \tag{3.5}$$

As an example, the $P = 0$ solution (neglecting ion motion) is $1 - \omega_{pe}^2/\omega^2 = 0$, or

$\omega = \omega_{pe}$, which is simply the plasma oscillation. The $\eta_\parallel^2 = R$ solution (neglecting ion motion) is the familiar R-wave

$$\eta_\parallel^2 = 1 - \frac{\omega_{pe}^2}{\omega(\omega - |\Omega_e|)},$$

$\eta_\parallel^2 = L$ is the L-wave, and the $\eta_\perp^2 = RL/S$ solution is the X-wave. The O-wave (again neglecting ion motion) is given by $\eta_\perp^2 = P$:

$$\eta_\perp^2 = \frac{k_\perp^2 c^2}{\omega^2} = 1 - \frac{\omega_{pe}^2}{\omega^2} \quad \Rightarrow \quad \omega^2 = \omega_{pe}^2 + k_\perp^2 c^2.$$

In the limit $\omega >> |\Omega_e|$ the X-wave matches the O-wave solution $\omega^2 = \omega_{pe}^2 + k_\perp^2 c^2$, while the R and L both become $\omega^2 = \omega_{pe}^2 + k_\parallel^2 c^2$; in fact the plasma in this limit becomes isotropic and is completely described by $\omega^2 = \omega_{pe}^2 + k^2 c^2$, as was first encountered in Section 1.4.2 with Equation (1.16).

3.1.2 Alfvén and Magnetosonic Waves

The plasma wave modes relevant for this chapter are in the ULF and low ELF range. Unlike the HF wave modes that only involve the "jiggling" of electrons, the wave period for the ULF/ELF range is so long that electrons respond essentially instantly and the ions are the ones doing the jiggling. In fact, "ultra" low means the wave frequency is much smaller than all ion gyrofrequencies in the system. In the notation from Appendix B.1 this translates to $\omega << \Omega_\alpha$, which for the ionosphere[4]

[4]Taking, say, the dominant O^+ ions in the F region of the ionosphere, the corresponding gyrofrequency is about 45 Hz, while those for O_2^+ or NO^+ is about half that. The ULF range by definition must have a frequency range far below these frequencies.

also means $\omega << \omega_{p\alpha}$. The ULF range dispersion relations can be extracted by taking the $\omega << |\Omega_\alpha|$ limit of the CPDR (B.15); this is the *magnetohydrodynamic* (MHD) limit, and the corresponding waves are referred to as MHD waves.

Taking a look back at the expression for R we see that taking the limit $\omega << \Omega_\alpha$ to first order in ω/Ω_α gives

$$R = 1 - \frac{1}{\omega}\sum_\alpha \frac{\omega_{p\alpha}^2}{\Omega_\alpha}\frac{1}{(1+\omega/\Omega_\alpha)} \simeq 1 - \frac{1}{\omega}\sum_\alpha \frac{\omega_{p\alpha}^2}{\Omega_\alpha}\left(1 - \frac{\omega}{\Omega_\alpha}\right) = 1 - \frac{1}{\omega}\sum_\alpha \frac{\omega_{p\alpha}^2}{\Omega_\alpha} + \sum_\alpha \frac{\omega_{p\alpha}^2}{\Omega_\alpha^2}. \tag{3.6}$$

The first sum on the RHS, $\sum_\alpha \omega_{p\alpha}^2/\Omega_\alpha$, is proportional to $\sum_\alpha q_\alpha n_\alpha$ and is thus zero by quasineutrality. The second sum on the RHS can be taken to be only over the ion species since the electron contribution to the sum is smaller by a factor of m_e/m_i. A similar treatment of L will give the same exact result (to first order in ω/Ω_α), thus we have

$$R, L \simeq 1 + \sum_i \frac{\omega_{pi}^2}{\Omega_i^2}, \tag{3.7}$$

where the sum is now over the ions only. S, D, and P simplify to

$$S \simeq R, L; \quad D \simeq 0; \quad P = 1 - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega^2} \simeq -\frac{\omega_{pe}^2}{\omega^2} \to -\infty. \tag{3.8}$$

Using (3.8) along with the dispersion relation Equation (B.15) and simplifying we obtain

$$\eta^2 = \frac{S^2 \sin^2\theta(1 \pm 1) + PS(1 + \cos^2\theta \mp \sin^2\theta)}{2(S\sin^2\theta + P\cos^2\theta)}. \tag{3.9}$$

Taking the lower sign first gives

$$\eta^2 = \frac{PS}{S\sin^2\theta + P\cos^2\theta} \Rightarrow \omega^2 = c^2\left(\frac{1}{S}k^2\cos^2\theta + \frac{1}{P}k^2\sin^2\theta\right) \simeq c^2\frac{1}{S}k^2\cos^2\theta. \tag{3.10}$$

Now, recall that S is given by

$$S = 1 + \sum_i \frac{\omega_{pi}^2}{\Omega_i^2} = 1 + \sum_i \frac{q_i^2 n_i}{\epsilon_0 m_i}\frac{m_i^2}{q_i^2 B_0^2} = 1 + c^2\frac{\mu_0\sum_i m_i n_i}{B_0^2} \equiv 1 + \frac{c^2}{v_A^2}, \tag{3.11}$$

where the quantity v_A, called the *Alfvén velocity*, is defined by

$$v_A = \sqrt{\frac{B_0^2}{\mu_0\rho_m}}, \tag{3.12}$$

and ρ_m is the total mass density of the plasma given by $\rho_m = \sum_i m_i n_i$. The dispersion relation now becomes

$$\omega^2 = \frac{c^2}{1 + c^2/v_A^2}k^2\cos^2\theta = \frac{v_A^2}{1 + v_A^2/c^2}k^2\cos^2\theta \simeq v_A^2 k^2\cos^2\theta, \tag{3.13}$$

where the approximation on the RHS is justified since $v_A << c$ in ionosphere. The wave mode characterized by the dispersion relation (3.13) is called the *Alfvén wave*, named after Hannes Alfvén who first theoretically predicted it. Expressing Equation (3.13) as $\omega = v_A k_z$ shows that the group velocity ($\mathbf{v}_g$) is directed only along the background magnetic field: $\mathbf{v}_g = v_A\hat{\mathbf{z}}$. Alfvén waves are associated with the *tension* stored in a magnetic field. Consequently, the propagation of Alfvén waves is similar to picturing the background field as being a taught string, and the "plucking" of this string as generating waves that travel along it. Considering that the charged

particles in a plasma are "stuck" on the field lines due to gyration, one could say that the field lines have an effective mass density of ρ_m. On the other hand, the Maxwell stress tensor says that field lines are under tension per unit area of B_0^2/μ_0. The speed of propagation on a string under tension T with linear mass density ρ_l is given by $v = \sqrt{T/\rho_l}$, thus giving an intuitive derivation of the Alfvén velocity:

$$v = \sqrt{\frac{T}{\rho_l}} = \sqrt{\frac{B_0^2}{\mu_0 \rho_m}} = v_A.$$

Since the magnetic field lines twist relative to one another but do not compress, the Alfvén wave is also commonly referred to as the *Shear Alfvén* (SA) wave.

Going back to Equation (3.9) and taking the top sign we obtain

$$\eta^2 = \frac{2S^2 \sin^2\theta + PS(1 - \sin^2\theta + \cos^2\theta)}{2(S\sin^2\theta + P\cos^2\theta)} = S, \tag{3.14}$$

which can be rewritten (as done for the SA wave) to give

$$\omega^2 = v_A^2 k^2. \tag{3.15}$$

This wave mode is the *MagnetoSonic* (MS) wave, and unlike the SA wave it *is* a compressional wave and propagates *isotropically*[5] rather than only along field lines; the group and phase speed of the MS wave are the same: $\mathbf{v}_p = \mathbf{v}_g = v_A\hat{\mathbf{k}}$. The MS wave, as its name suggests, resembles sound waves since the restoring force of the wave is the *magnetic pressure* rather than magnetic tension. In the case of a *warm*

[5]There is a slight caveat: for the case of a point source, the generated MS waves propagate *almost* isotropically. Namely, they propagate everywhere *except* exactly along the background field, since field lines do not compress in the parallel direction.

plasma, the pressure term would have been kept and the "sonic nature" would be more explicit since the restoring force would then originate from both the thermal and magnetic pressures. Incidentally, there is a dimensionless parameter called the plasma-β (or just β) that is defined to be a ratio of the total thermal pressure ($p = p_e + p_i$) to the magnetic pressure ($B^2/2\mu_0$):

$$\beta \equiv \frac{2p\mu_0}{B^2}. \tag{3.16}$$

Neglecting the pressure term is equivalent to assuming a low-β plasma ($\beta << 1$), which is often true for the ionosphere. Figure 3.1 schematically shows the SA wave (left) and MS wave (right), where the characteristics described above (i.e. tension and compression) can be seen.

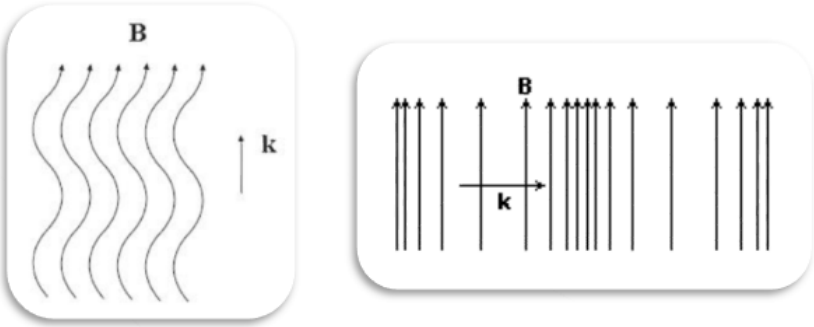

Figure 3.1: Schematic showing the SA wave (left) and MS wave (right); the tension (but no compression) nature of the SA wave is clearly seen, while the opposite is seen for the MS wave.

The SA and MS waves can be derived from a different approach as well. Rather than taking limits of the general CPDR to find the SA/MS wave dispersion relations, it is possible to start with the basic equations and make the necessary approximations from the start. The simplified set of equations that emerge are the single-fluid MHD equations. The details will not be reproduced here for brevity and since it

is carried out in many standard introductory plasma physics textbooks (e.g. [12]), but a brief derivation of the evolution equations for the fields in the MHD limit is given in Appendix B.2 in the form of Equations (B.21); it is shown that by assuming one (singly ionized) ion species, a constant (vertical) background magnetic field $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$, constant background plasma density n_0, no collisions, and a low-β plasma, the time evolution of the electric field and vector potential in the MHD limit is given by

$$\frac{\partial \mathbf{E}}{\partial t} = -v_A^2 \left[(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{A}) \times \hat{\mathbf{z}}\right] \times \hat{\mathbf{z}}, \tag{3.17a}$$

$$\frac{\partial \mathbf{A}}{\partial t} = -\mathbf{E}. \tag{3.17b}$$

The fields $\mathbf{E}$ and $\mathbf{A}$ can be solved numerically once the proper source of SA or MS waves is introduced into the RHS of (3.17a). However, for simplicity the sources will be left out since that will be discussed later in the chapter. By assuming a 2D geometry the set of Equations (3.17a) and (3.17b) can be solved numerically on the computer with a number of numerical recipes. We can arbitrarily choose the simulation domain to be in the x-z plane, thus making all partial y derivatives zero in the set of equations. For the problem at hand a spectral method in space and a fourth order Runge-Kutta (RK4) in time is chosen. When the SA mode is excited in Equations (3.17a) and (3.17b) the result is a non-zero B_y and E_x, while the remaining components of the fields are zero, as is illustrated by Figure 3.2a,b. As expected, the SA wave propagated only vertically, i.e. along the background magnetic field. The color scale in the figure is arbitrary; the purpose of the figures is to simply show the characteristic wave behavior, not to compute amplitudes. The axes are normalized

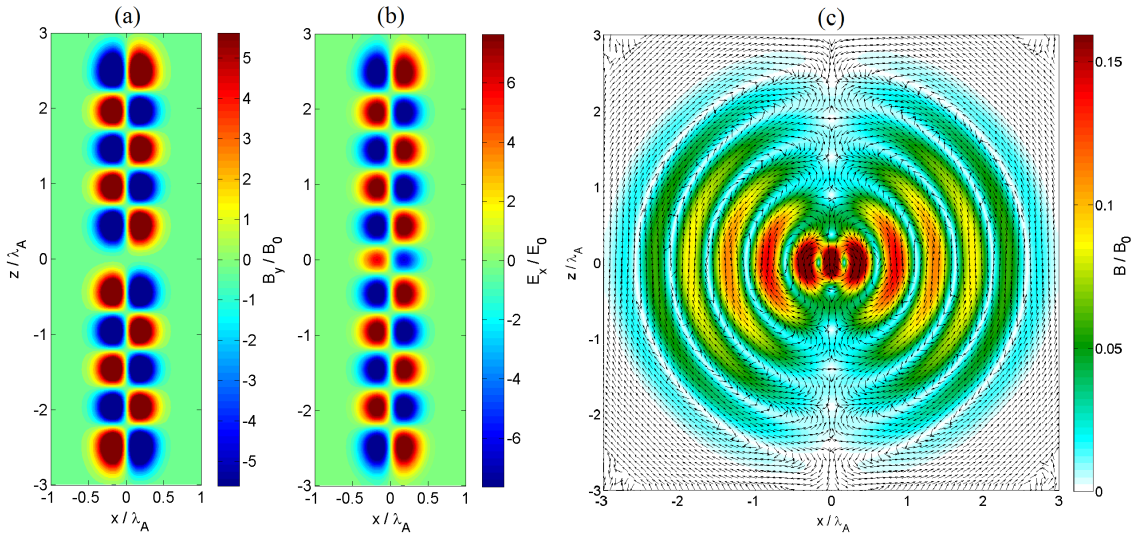

Figure 3.2: Simulation of SA waves, showing the only non-zero components B_y (a) and E_x (b), and MS waves (c) showing the **B** vector field superimposed on the magnitude $|\mathbf{B}|$. Axes are measured relative to λ_A, while the amplitudes are arbitrary.

to the Alfvén wavelength (λ_A) and time (not shown) is normalized to the Alfvén period (τ_A), which are related to each other by $v_A = \lambda_A/\tau_A$.

Exciting the MS mode results in a rather different wave pattern, as is shown by Figure 3.2c. The figure reveals the magnetic vector field superimposed on the magnitude of the field, $|\mathbf{B}|$. The magnetic field is entirely in the x-z plane ($B_y = 0$), which means the electric field (not shown) only has an E_y component, with an intensity pattern similar to $|\mathbf{B}|$.

It is also possible to solve Equations (3.17) for a 3D geometry in a very similar way. Choosing again a spectral method in space and stepping everything forward in time with the RK4, a more intuitive illustration of the SA and MS waves can be obtained. Figure 3.3 shows the results of SA wave excitation in a 3D geometry; slices of the field amplitudes along with vector fields superimposed are shown. The magnetic field (left) only has an azimuthal component, with an intensity pattern similar to the figure above. On the other hand, the electric field (right) is purely

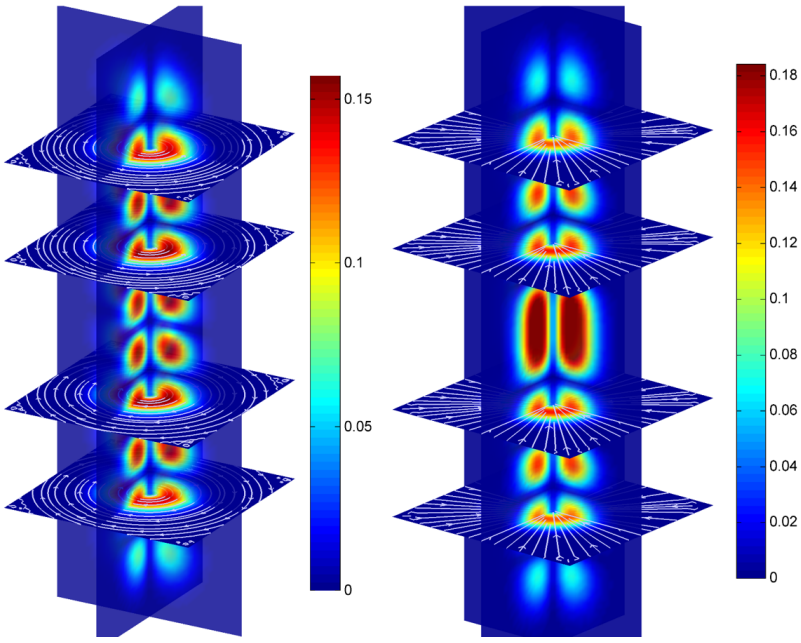

Figure 3.3: Propagation of shear Alfvén waves in 3D, showing slices of the magnetic (left) and electric (right) vector fields (white arrows) superimposed on their magnitudes (color plots).

radial, as one would expect from $\mathbf{E}$ and $\mathbf{B}$ orthogonality. It is quite clear from the figure that shear Alfvén waves can be characterized by a magnetic field with non-zero perpendicular curl ($\boldsymbol{\nabla}_\perp \times \mathbf{B} \neq 0$) and an electric field with a non-zero perpendicular divergence ($\boldsymbol{\nabla}_\perp \cdot \mathbf{E} \neq 0$).

Figure 3.4 shows the results of MS wave excitation in a 3D geometry, with slices of the field amplitudes and vector fields similarly shown. The magnetic field slices (left) show an intensity and vector field pattern very similar to the 2D version, and has no azimuthal component. On the other hand, the electric field (right) is purely azimuthal, as it should be. These 3D fields are really no different than rotating the 2D cases about the magnetic field, and having the fields drop off as $1/r$ (3D) rather than $\sim 1/\sqrt{r}$ (2D); although, no change in amplitude drop-off occurs for SA waves when making the transition from 2D to 3D, since they always propagate in one direction. It is clear from Figure 3.4 that MS waves can be characterized

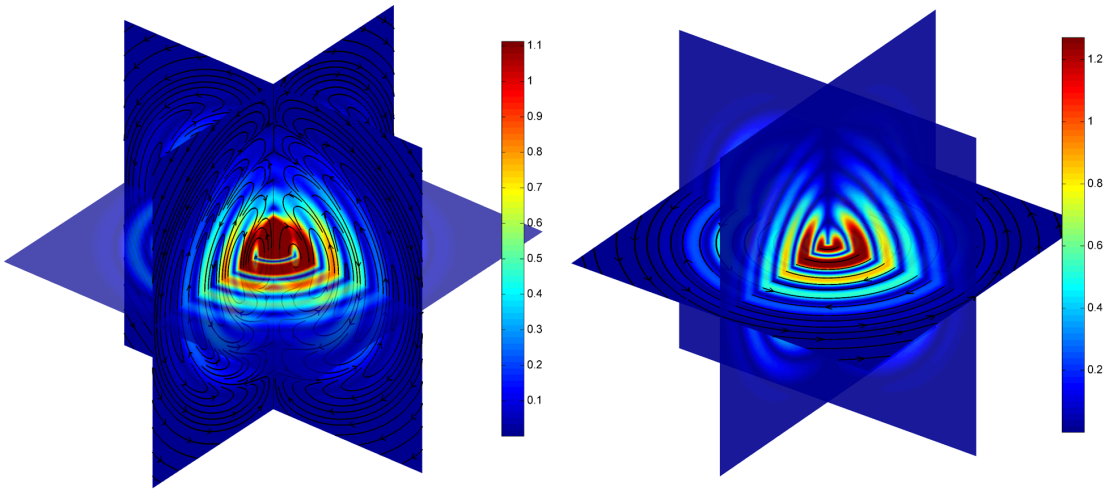

Figure 3.4: Propagation of magnetosonic waves in 3D, showing slices of the magnetic (left) and electric (right) vector fields (black arrows) superimposed on their magnitudes (color plots).

by a non-zero perpendicular curl of the electric field ($\boldsymbol{\nabla}_\perp \times \mathbf{E} \neq 0$). It turns out they are also characterized by the vertical magnetic field component B_z, but that is less obvious to understand. Note that there is a clear correspondence for the field components when going from 2D to 3D: $\hat{\mathbf{x}} \leftrightarrow \hat{\mathbf{r}}$, $\hat{\mathbf{y}} \leftrightarrow \hat{\boldsymbol{\phi}}$, $\hat{\mathbf{z}} \leftrightarrow \hat{\mathbf{z}}$, where the unit vectors $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\phi}}$, and $\hat{\mathbf{z}}$ refer to the cylindrical coordinate system.

3.2 Modulated Heating: Using the Ionosphere as an Antenna

A significant amount of research has been put into ways of generating MHD waves in the ionosphere. The main motivational factors have come from two possible practical applications. The first application is for RBR in order to inhibiting the damage done to satellites by high energy plasma particles that always populate certain regions of the magnetosphere. These regions are called the *radiation belts* (or Van Allen radiation belts), and a sudden surge of energetic particles from abnor-

mally high levels of solar activity[6] could result in the destruction of communication and navigation systems on-board satellites, and thus deal a serious blow to our infrastructure and national security. The generation of Alfvén waves and subsequent injection into the radiation belts could be an efficient method for precipitating the harmful high energy ions via resonant pitch angle scattering [51], and hence reducing their harmful impact on satellite electronics. As mentioned in Chapter 2, VLF wave injection can similarly precipitate harmful electrons out of the radiation belts.

The fact that MHD waves are of such low frequency has caught the attention of the Navy, and accounts for the second application. Conventional radio communication cannot establish a connection with submerged submarines due to the exponential decay of EM waves inside the electrically conductive seawater, thus requiring submarines to rise the surface before communication can be achieved. The characteristic length scale of this exponentially decay - the *skin depth* - increases as the frequency of the EM wave decreases. If ULF or ELF waves could be generated, then communication (albeit at a very low bandwidth) could be maintained with submarines at much lower depths, which is favorable from a military standpoint. However, generating ULF/ELF requires an enormously large antenna; so much so that constructing such an antenna is very difficult and costly from an engineering point of view (for ELF), or outright impossible (for ULF). If fact, ELF facilities were constructed by the Navy and operated in the range of 20 - 40 Hz for two decades, but eventually decommissioned in the early 2000s. These antennas were tens of miles long, and took considerable power to operate. Similar facilities with better perfor-

[6] Another possibility in addition to abnormally high levels of solar activity is the deliberate high altitude detonation of a nuclear bomb [35].

mance had been conceived but never constructed, since they would likely require several thousands of miles of buried cables, several hundred megawatts of power, and had been projected to cost billions (e.g. Project Sanguine).

A possible resolution to ELF generation came from ionospheric physics. It was discovered that the D and E regions of the ionosphere (at certain latitudes) have a naturally occurring current system that is being driven by the sun's solar wind flowing past the earth's magnetosphere. This current system, termed the *electrojet*, can be understood crudely by considering the ideal Ohm's law (B.18). If there exists an electric field, the plasma will flow in such a way that Equation (B.18) is satisfied. The opposite is true as well: if plasma flows past the earth's magnetic field $\mathbf{B}_0$ (as the streaming solar wind does), then it will set up an electric field that satisfies $\mathbf{E} = -\mathbf{v} \times \mathbf{B}_0$. This amounts to a potential difference being set up between the earth's field lines, which electrons will quickly tend to neutralize. However, electrons cannot easily transport across field lines, since field lines in the collisionless magnetosphere are only conductive parallel to the magnetic field (and in this sense act as equipotential "wires"). Instead, electrons flow down the field lines toward the earth and "complete the circuit" through the conductive ionosphere. If fact, it is the ionosphere's *Hall conductivity* that allows the current closure. Moreover, it was a known fact that electron heating of the D/E region plasma results in an increased Hall conductivity; therefore *modulated heating* of the D/E region ionosphere in the presence of the electrojet sets up an alternating current, effectively creating an antenna in the D/E region that radiates at the frequency of the modulation.

The necessity of an electrojet for this method of EM wave generation is obvi-

ously important. Fortunately there are two regions on the earth where an electrojet can exist. The equatorial region is one of these regions and tends to have the strongest electrojet, but unfortunately has no ionospheric heaters to date. The other is the high latitude polar region, which includes the locations of the HAARP and EISCAT heaters, and has been termed the Polar ElectroJet (PEJ). While not as stable as the equatorial electrojet, PEJ modulation has been successfully used by several scientists at the HAARP facility to generate EM waves in the ULF/ELF/VLF range [46, 4, 36, 38, 40, 41, 45, 44, 43, 13]. The antenna associated with PEJ modulation - the PEJ antenna - injects ELF/VLF waves in the earth-ionosphere wave guide, as well as whistler and shear Alfvén waves in the magnetosphere and the radiation belts. Wave generation from ULF frequencies as low as mHz to VLF frequencies up to 10-20 kHz has been confirmed by both ground and satellite observations. ELF generation by electrojet modulation turns out to be much less costly than the Navy VLF transmitters used for submarine communication. The reason is simple: it takes a huge amount of power for a sufficiently large current to flow through the large (tens of miles sized) ELF antennas, since the current closes through the ground (so-called ground dipole antenna). The electrojet currents (if available) are already given to us "for free". Of course power is still necessary to make an electrojet antenna, but the power is fed into the HF transmitter instead, which is significantly more efficient than an ELF transmitter. Moreover, the PEJ has the possibility of going to much lower frequencies without much extra effort; the only added difficultly is running the heater for a longer time due to the long modulation period.

3.2.1 Ionospheric Current Drive (ICD)

The major drawback of the PEJ antenna is the requirement of strong electrojet currents that render the process inapplicable in mid-latitude regions and with spotty reliability in the polar region. Another method for generating low frequency waves was developed that does not rely on the presence of electrojet currents. This process, known as Ionospheric Current Drive (ICD), relies on modulated heating of the F region. Rather than waiting for electrojet currents, this method effectively *drives* its own currents (hence the name) in the E region in the following way: first F region modulated heating creates a modulated electron pressure gradient in the heated region. An increased electron pressure forces electrons to drift in a direction perpendicular to both the gradient and the background magnetic field, generating a *diamagnetic current*. The effect is essentially to produce an oscillating magnetic dipole antenna in the form of an oscillating diamagnetic current in the F region. It is referred to as diamagnetic since the current flows in a direction such that the magnetic field *it* generates is in the opposite direction of the background field (Lenz's law). The end result is a *depression* of the magnetic field magnitude, without twisting of the field. Therefore the magnetic pressure is modulated, and so it comes as no surprise that *magnetosonic* waves are generated at the frequency of modulation. Some of the MS waves then travel down toward the E region and encounter the Hall conductivity, where the electric field from the MS wave *drives oscillating Hall currents*. These oscillating currents act as a *secondary antenna* that then radiate SA waves, some of which penetrate toward the ground and get injected in the earth-ionosphere waveguide as conventional EM waves. This whole process

is scematically illustrated by Figure 3.5. Due to the slow response time of the F

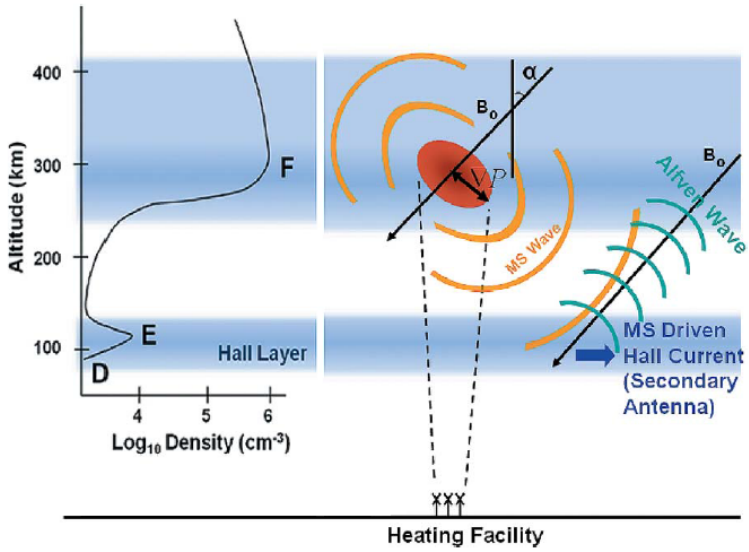

Figure 3.5: Schematic illustrating the ICD concept. MS waves (orange) are generated by modulated F region heating (red ellipse), which then propagate down to the E region and generate SA waves (cyan) by driving Hall currents. (Reproduced from Figure 1 of Papadopoulos [2011a].)

region the operation of the ICD antenna is limited to frequencies below 50-60 Hz (ULF and ELF). Experiments carried out at HAARP indicate that in addition to ELF waves injected in the earth-ionosphere waveguide, ICD injects SA waves in the magnetosphere along the MZ of the heater and MS waves in the Alfvénic duct[7] [39, 37, 14].

3.2.2 Mathematical Model of ICD

A theoretical description of the ICD concept requires a mathematical model that contains the physics described above. Two different approaches to modeling ICD were investigated, both of which used a cold plasma approach and derived

[7]This is entirely different from the ducts in Chapter 2. The Alfvénic duct simply refers to the waveguide created by the E region and topside F region; MS waves get trapped in this altitude range since the Alfvén speed has a minimum at the F_2 peak but rises considerably for altitudes much lower or higher, thus leading to their total internal reflection.

evolution equations for the fields. The first model, which is based on [29], demonstrated the ICD concept [39]. This model uses a cylindrical 2D geometry to study MS wave generation by F region heating and subsequent SA wave coupling via the Hall conductivity. The model assumes an exactly vertical geomagnetic field, which is approximately true for polar region heaters such as HAARP. The MHD limit $\omega << \omega_{ci}$[8] is assumed as well, making the model valid only for ULF range waves. The derivation of the model equations will not be reproduced but will be quoted instead, along with a brief discussion. The dynamical variables are defined similarly to [29]:

$$Q = \boldsymbol{\nabla}_\perp \cdot \mathbf{E}_\perp, \quad M = (\boldsymbol{\nabla}_\perp \times \mathbf{E}_\perp) \cdot \hat{\mathbf{z}}, \quad J_z = (\boldsymbol{\nabla}_\perp \times \mathbf{B}_\perp) \cdot \hat{\mathbf{z}}.$$

With the addition of a localized source due to the pressure gradient driven by HF heating, the equations of the model can be derived to be [39]

$$\begin{aligned}
&\left(\epsilon\frac{\partial}{\partial t} + \sigma_P\right) Q = -\sigma_H M - \frac{\partial J_z}{\partial z}, \\
&\left(\epsilon\frac{\partial}{\partial t} + \sigma_P\right) M = \sigma_H Q - \frac{1}{\mu_0}\nabla^2 B_z + \frac{1}{B_0}\nabla_\perp^2 \delta p_\perp, \\
&\frac{\partial B_z}{\partial t} = -M, \quad \mu_0\frac{\partial J_z}{\partial t} = -\frac{\partial Q}{\partial z} + \boldsymbol{\nabla}_\perp^2 E_z, \quad \left(\epsilon_0\frac{\partial}{\partial t} + \sigma_\parallel\right) E_z = J_z.
\end{aligned} \tag{3.18}$$

Here σ_H, σ_P, and $\sigma_\parallel$ are the Hall, Pedersen, and parallel conductivities, and $\epsilon(z)$ is the plasma dielectric function given by

$$\epsilon(z) = \frac{c^2}{v_A^2(z)\left[1 + \nu_{in}^2(z)/\Omega_i^2\right]}. \tag{3.19}$$

[8]Here ω_{ci} is the ion gyrofrequency and will be used interchangeably with the Ω_i; the notation that is consistent with the relevant literature will take precedence.

In Equation (3.19) $v_A(z)$ is the (altitude dependent) Alfvén speed, $\nu_{in}(z)$ the effective ion-neutral collision frequency and Ω_i is the ion gyrofrequency. The pressure source due to HF heating is assumed to be of the form

$$\delta p_\perp(r,z,t) = n(z)k_B T_{heat} \tanh^2(t/D_t)\left[1+\cos(\omega t)\right]\exp\left[-\frac{r^2}{D_r^2}-\frac{(z-z_{heat})^2}{D_z^2}\right]. \tag{3.20}$$

This term describes F region heating near the F_2 peak located at z_{heat} with average transverse electron temperature T_{heat} over the heated region, assumed to have Gaussian profiles with widths D_r and D_z in the radial and z directions, respectively. The transient temporal behavior of the source is described as a smooth ramp-up with characteristic time D_t, while also varying harmonically in time with (angular) frequency $\omega = 2\pi f$. The equations are solved numerically over the atmosphere-ionosphere domain using a discrete Fourier-Bessel transform in the r coordinate and the Crank-Nicholson scheme for z and time t. The use of the Crank-Nicholson implicit scheme allows for the solution in the atmosphere with large time steps without the possibility of numerical instability; the Alfvén time scale in the ionosphere is, of course, resolved. As the code steps forward in time, it periodically recovers the perpendicular field components using Q, M, and J_z, and saves the data. Simulating the above with $z_{heat} = 300$ km, $D_t = .25$ sec, $D_r = 100$ km, $D_z = 20$ km, $T_{heat} = 5000$ K, and a modulation frequency of $f = 10$ Hz gives the results shown in Figure 2 of [39], reproduced here in Figure 3.6. The simulation results show the generation of MS waves (a) for two different times. The MS waves quickly propagate down and drive Hall currents, which then radiate SA waves (b). Note that due to the increased phase speed at higher altitudes, the MS waves refract near the top of the

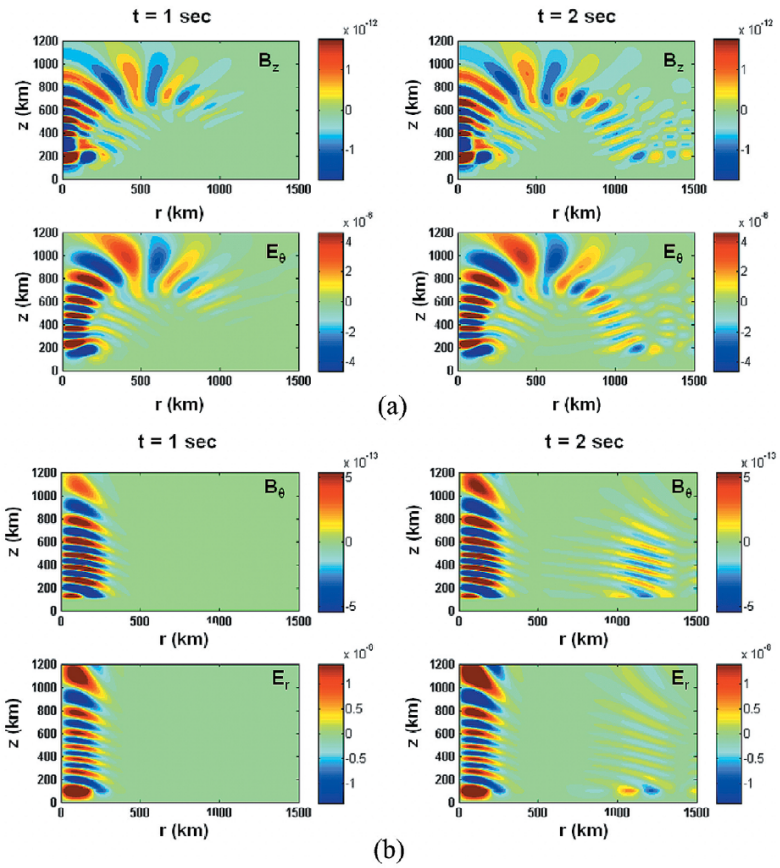

Figure 3.6: Results of simulating ICD using the model equations. Generation of MS waves (a) can be seen for two different times ($t = 1, 2$ sec), along with the generation of SA waves (b) by the subsequent MS-driven Hall current; **E** and **B** are in SI units. (Reproduced from Figure 2 of [39].)

domain and travels back downward (Figure 3.6a, $t = 2$ sec), eventually interacting with the E region again and driving SA waves near $r = 1200$ km (Figure 3.6b, $t = 2$ sec). Soon after the first simulation studies of ICD were published, a summary of some proof of principle experiments were demonstrated and published in Reference [37]. F region modulated heating experiments during *underdense* conditions were conducted over the course of several hours and showed magnetometer measurements near HAARP registering signals at the modulation frequency. Several frequencies ranging from 0.2 Hz - 70 Hz were used and resulted in the ground detection of magnetic field amplitudes in the range 0.05 pT - 1 pT. These results are reproduced

in Figure 3.7, which shows magnetic field amplitudes observed on the ground near Gakona, Alaska, for an experiment conducted on 09/09/2009 (a) and on 11/03/2011 (b). The magnetometer registered field values ranging from 0.05 - 0.4 pT in the former case and 0.2 - 1 pT in the latter case. The amplitudes seem to rise from the

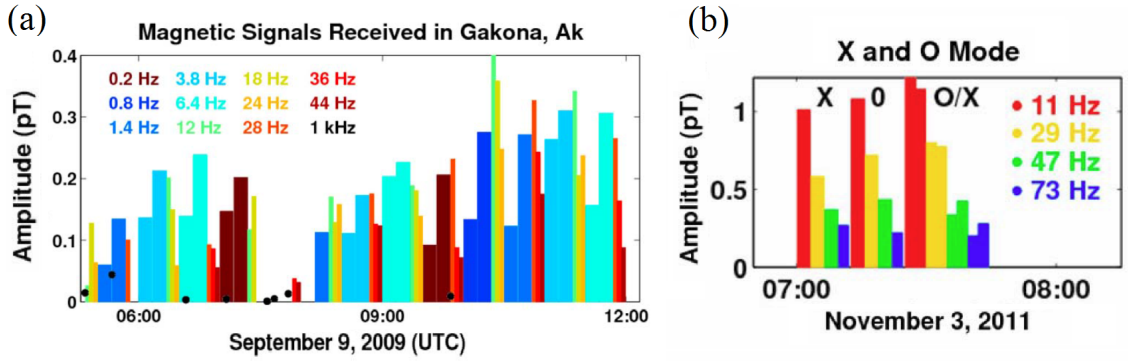

Figure 3.7: Proof of concept experimental results of the ICD method. The first experiment (a) shows ground field amplitudes for a wide range of frequencies, showing a peak at 12 Hz; and similarly for another experiment (b), but only for 11 Hz and above , showing a $1/f$-like behavior. (Adapted from Figures 2 and 4 of [37].)

lowest frequencies to 12 Hz (Figure 3.7a), after which there is $1/f$ drop-off behavior all the way to the highest successful ICD frequency near 70 Hz (Figure 3.7b). The $1/f$ behavior indicated that the F region heating/relaxation time scale had been reached. It was shown that, at least during underdense conditions, O-mode heating and X-mode heating gave very similar results. Moreover, the power scaling of ICD was demonstrated to be linear, since HAARP operating at 50% power resulted in ground signatures that were approximately half the amplitude achieved when HAARP operated at 100% power. Further discussion of the experimental results and subsequent comparisons with modeling is the subject of later sections.

To study ICD in more general background field geometries and not have the MHD restriction ($\omega << \omega_{ci}$), Reference [14] developed an ICD model with more flexibility. The model is valid for the range $\omega << \omega_{ce}$ and thus accommodates

ULF, ELF, and VLF wave modes such as the SA, MS, and whistler waves. The assumptions of the model are as follows:

- Small amplitude waves (linearized equations)
- Static (in time), but otherwise arbitrary geomagnetic field, $\mathbf{B}_0$
- Electron inertia is negligible (valid for $\omega << \omega_{ce}$)
- Static, but otherwise arbitrary background plasma density, n_0
- Single ion species (valid for most of the F region), so $n_i = n_e \equiv n_0$
- Displacement current is negligible (valid for ULF/ELF frequencies)
- HF heating is described by a localized electron temperature perturbation, T_e

The complete set of necessary equations includes Faraday's and Ampère's laws, and the electron and ion momentum equations that abide by the above assumptions:

$$\frac{\partial \mathbf{B}}{\partial t} = -\boldsymbol{\nabla} \times \mathbf{E}, \tag{3.21a}$$

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 e n_0 (\mathbf{v}_i - \mathbf{v}_e), \tag{3.21b}$$

$$0 = -\frac{e}{m_e}(\mathbf{E} + \mathbf{v}_e \times \mathbf{B}_0) - \nu_{en}\mathbf{v}_e - \frac{\boldsymbol{\nabla} p_e}{m_e n_0}, \tag{3.21c}$$

$$\frac{\partial \mathbf{v}_i}{\partial t} = \frac{e}{m_i}(\mathbf{E} + \mathbf{v}_i \times \mathbf{B}_0) - \nu_{in}\mathbf{v}_i, \tag{3.21d}$$

where $p_e = n_0 k_B T_e$ is the external pressure source due to the electron temperature perturbation, and ν_{en} (ν_{in}) is the electron-neutral (ion-neutral) collision frequency. It would be convenient to have a set of decoupled equations for the fields alone, which can be achieved by various combinations of time differentiation and substitution.

The details of this are omitted here, but included in Appendix C. The result is a set of coupled time evolution equations for the electric field and vector potential:

$$\frac{\partial \mathbf{A}}{\partial t} = -\mathbf{E}, \tag{3.22a}$$

$$\frac{\partial \mathbf{E}}{\partial t} = -\omega_{ci}(\Gamma_{en} + \Gamma_{in})\mathbf{E} + \frac{\overline{\overline{\boldsymbol{\epsilon}}}^{-1}\left[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{A}\right]}{\mu_0} - \frac{\overline{\overline{\mathbf{R}}}_e\left[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}\right]}{\mu_0 \tilde{\sigma}} + \left(\omega_{ci}\overline{\overline{\mathbf{R}}}_i - \frac{\partial}{\partial\, t}\right)\frac{\boldsymbol{\nabla} p_e}{e n_0}, \tag{3.22b}$$

where the *dielectric tensor* $\overline{\overline{\boldsymbol{\epsilon}}}$, tensors $\overline{\overline{\mathbf{R}}}_e$ and $\overline{\overline{\mathbf{R}}}_i$, the *conductivity tensor* $\overline{\overline{\boldsymbol{\sigma}}}$ (not seen here), and other quantities are defined in Appendix C.

What remains is to specify the HF heating source, which we will do in a similar way to Equation (3.20). Namely, a localized electron pressure perturbation is assumed:

$$\boldsymbol{\nabla} p_e = k_B \boldsymbol{\nabla} n_0 T_e \simeq n_0 k_B \boldsymbol{\nabla} T_e, \tag{3.23}$$

where the gradient of the temperature perturbation is taken to be

$$\boldsymbol{\nabla} T_e = T_{heat} \tanh^2(t/D_t)\cos(\omega t)\boldsymbol{\nabla} \exp\left[-\frac{x^2}{D_x^2} - \frac{(z - z_{heat})^2}{D_z^2}\right]. \tag{3.24}$$

T_{heat} is the modulation amplitude of the electron temperature, D_t is the rise time, D_x and D_z are the widths of the heated region in the x and z directions, z_{heat} is the heating altitude, and ω is the HF heating modulation frequency[9]. Equation (3.23) assumes that the spatial variation in the density is negligible. This is valid near the F_2 peak where the density profile is nearly flat; the relative contribution of the

[9]Note that here we have chosen the time dependence as $\cos(\omega t)$, rather than $1 + \cos(\omega t)$ like in Equation (3.20). It turns out that a purely harmonic time dependence is advantageous for subsequent Fourier analysis of the ICD generated waves. The constant factor 1 introduces a transient DC component in the spectrum, which is of no use to us.

density gradient is explored later in the chapter.

Note that in the limit $\Gamma_{en} = \Gamma_{in} = 0$, and for $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$, the electric field equation from (3.22) reduces to

$$\frac{\partial \mathbf{E}}{\partial t} = -v_A^2 \left[(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{A}) \times \hat{\mathbf{z}}\right] \times \hat{\mathbf{z}} - \frac{c^2 \omega_{ce}}{\omega_{pe}^2} \left[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}\right] \times \hat{\mathbf{z}} + \omega_{ci} \frac{\boldsymbol{\nabla} p_e}{e n_0} \times \hat{\mathbf{z}} - \frac{\partial}{\partial t} \frac{\boldsymbol{\nabla} p_e}{e n_0}, \tag{3.25}$$

which is the collisionless Hall-MHD equation (with the HF heating sources). For frequencies much lower than the ion gyrofrequency the second term on the RHS can be neglected, as well as the second source term. The full system in this limit becomes

$$\frac{\partial \mathbf{E}}{\partial t} = -v_A^2 \left[(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{A}) \times \hat{\mathbf{z}}\right] \times \hat{\mathbf{z}} + \omega_{ci} \frac{\boldsymbol{\nabla} p_e}{e n_0} \times \hat{\mathbf{z}}, \tag{3.26a}$$

$$\frac{\partial \mathbf{A}}{\partial t} = -\mathbf{E}, \tag{3.26b}$$

which is exactly the ideal MHD equations (3.17) from earlier, only this time with a source of waves. This source will only generate MS waves, and is in fact the source used to obtain the results of Figures 3.2c and 3.4.

Consistency requires that in the $\omega << \omega_{ci}$ limit we should recover wave dynamics similar to the ICD model equations (3.18) used by Reference [39]. Indeed, in this limit, the set of Equations (3.22) give very similar results to Equations (3.18) because simulations (not shown) performed by the author have show agreement between the two models to within about 15%.

3.3 Results of ICD modeling

The ICD model described by Equations (3.22) can be solved numerically for a plasma density $n(\mathbf{r})$ and background field $\mathbf{B}_0(\mathbf{r})$ of one's choosing. The region of primary interest is within ± 100 km along the ground (directly below the source), because ground magnetometer measurements are relatively close to HAARP. Since the ionosphere has little variation in the horizontal direction over ~ 100 km, the density will be taken to vary only in altitude for simplicity: $n = n(z)$. The background field for most cases to be discussed here is geared towards studying HAARP ionospheric heating, and to that end the field lines are nearly vertical. Some simulation runs are conducted with field lines tilted at 14° as at HAARP, though several other runs with purely vertical field lines will be presented as well. Even though tilted field lines are more accurate for HAARP, the motivation behind using a purely vertical $\mathbf{B}_0$ is that it facilitates the interpretation of simulation results. The focus is on propagation towards the ground and topside ionosphere, thus we assume the magnitude $|\mathbf{B_0}| = B_0$ to be a constant. For numerically solving the Equations (3.22), the z-derivatives are approximated by second-order centered finite differences, x-derivatives are approximated by the pseudo-spectral method, and the system is stepped forward in time with the RK4 method [14]. Atmospheric electric and magnetic fields are assumed to be quasi-static and solved separately using analytic methods for altitudes below approximately 90 km, and then matched to the ionospheric fields with appropriate boundary conditions; outflow boundary conditions are used for the top of the simulation domain [14].

Ionospheric profiles are chosen to have a form similar to those given in Ref-

erence [14]. For example, the plasma density is modeled by a simple Chapman profile:

$$n_0(z) = n_{max} \exp(1 - \xi - e^{-\xi}), \quad \xi = \frac{z - z_{max}}{H}, \tag{3.27}$$

where n_{max} is the peak plasma density at the F_2 peak located at the altitude z_{max}, and H is the ionospheric length scale. The altitude dependence of the normalized electron and ion collision frequencies $\Gamma_{en} = \nu_{en}/\omega_{ce}$ and $\Gamma_{in} = \nu_{in}/\omega_{ci}$ are given by

$$\Gamma_{en}(z) = \Gamma_{en0} \left[\exp\left(\frac{z_{en0} - z}{H_{en0}} \right) + \exp\left(\frac{z_{en1} - z}{H_{en1}} \right) \right], \tag{3.28}$$

and

$$\Gamma_{in}(z) = \Gamma_{in0} \left[\exp\left(\frac{z_{in0} - z}{H_{in0}} \right) + \exp\left(\frac{z_{in1} - z}{H_{in1}} \right) \right]. \tag{3.29}$$

To run simulations, we set the following typical values for nighttime ionospheric parameters: $n_{max} = 4.9 \times 10^4$ cm^{-3} (corresponding to $f_0F_2 = 2$ MHz), $z_{max} = h_mF_2 = 275$ km, $H = 97$ km, $B_0 = 4 \times 10^{-5}$ T (corresponding to $f_{ci} = 38$ Hz), $\omega_{ce} = 7.0 \times 10^6$ s^{-1} ($f_{ce} = 1.1$ MHz). For the collision frequency profiles we set the corresponding parameters as follows: $\Gamma_{en0} = 10^{-4}$, $z_{en0} = 110$ km; $H_{en0} = 20$ km, $z_{en1} = 110$ km, $H_{en1} = 4.35$ km, $\Gamma_{in0} = 0.5$, $z_{in0} = 120$ km, $H_{in0} = 20$ km, $z_{in1} = 120$ km, $H_{in1} = 5$ km. With these parameters we can plot various quantities, such as $n_0(z)$, $\Gamma_{en}(z)$, $\Gamma_{in}(z)$, $v_A(z)$, and the conductivities σ_P, σ_H, and $\sigma_\parallel$, as is illustrated by Figure 3.8. The various ionospheric profiles are plotted against altitude (in km). The top panels, from left to right, are the plasma density, normalized electron/ion collision frequencies $\Gamma_{en}(z)/\Gamma_{in}(z)$, and close-ups of $\Gamma_{en}(z)/\Gamma_{in}(z)$ showing their dependence in the E region and bottomside F region. The bottom panels, from left to

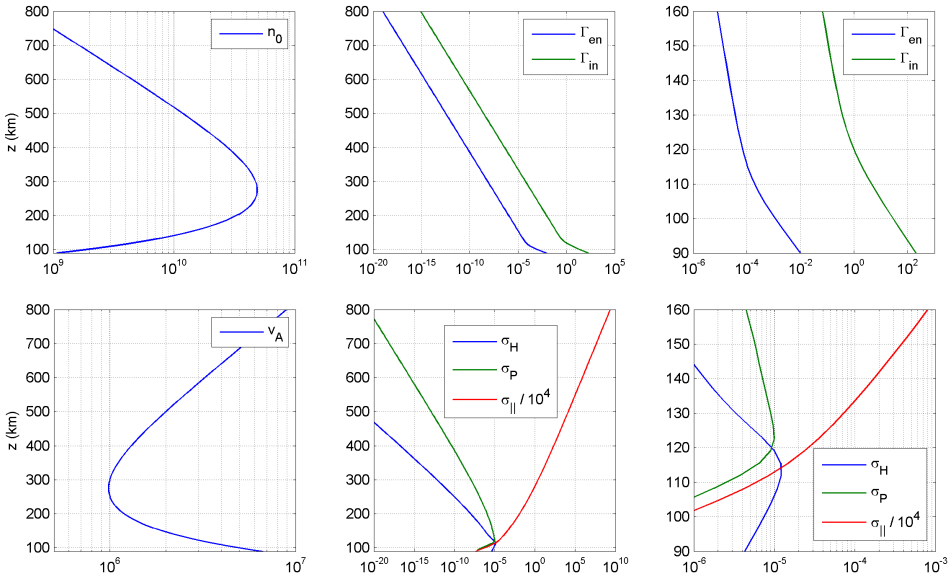

Figure 3.8: Various ionospheric profiles for a typical nighttime ionosphere, plotted against altitude z in km. The top panels, from left to right, are the plasma density, normalized electron/ion collision frequencies, and their close-ups. The bottom panels, from left to right, are the Alfvén speed, Pedersen/Hall/parallel conductivities, and their close-ups; all quantities are in SI units.

right, are the Alfvén speed, Pedersen/Hall/parallel conductivities, and a close-up of these conductivities. All quantities plotted on the x-axis are in SI units. It can be seen from the $\Gamma_{in}(z)$ profile that the ions are magnetized for altitudes above 120 km ($\Gamma_{in} < 1$) but loose their magnetization ($\Gamma_{in} > 1$) below 120 km, which essentially signifies the start of the E region. The electrons, on the other hand, are magnetized in the entire simulation domain ($\Gamma_{en}(z) < 1$) since they become demagnetized at much lower altitudes: closer to 65 km. Note also that as the ions become demagnetized, the Hall conductivity rises rapidly; this was mentioned and explained in Section 1.3.2.

For the heating source, we choose small vertical and horizontal widths for simplicity: $D_x = D_z = 10$ km. We also let $T_{heat} = 1000$ K (which can be readily attained with the HAARP heater), set the modulation frequency to $f = 5$ Hz, and

assume a purely vertical background field. With all of these inputs, the ICD model (3.22) can be numerically computed to give the field distributions shown by Figure 3.9. The MS wave components are the left three panels, showing B_x, B_z[10], and E_y.

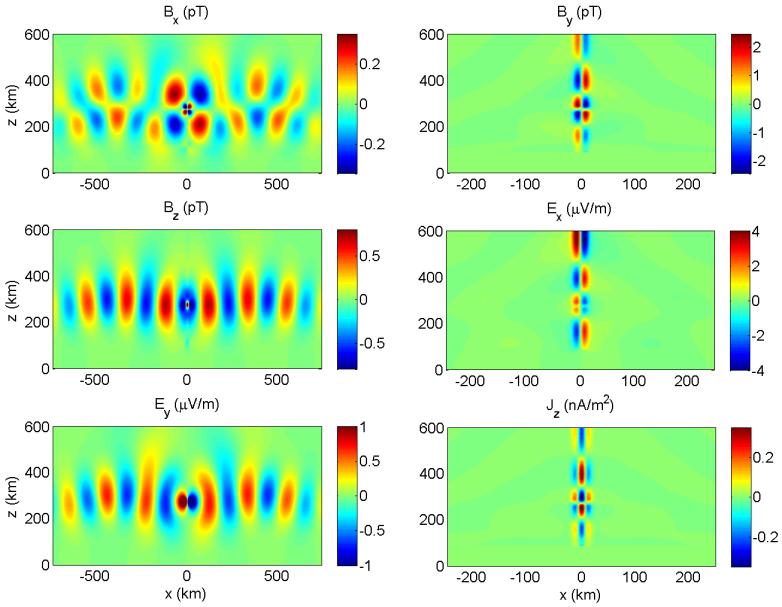

Figure 3.9: Simulation results showing the MS components (left panels), corresponding to B_x, B_z, and E_y and SA components (right panels), corresponding to B_y, E_x, and J_z.

The electric field $E_y\hat{\mathbf{y}}$ and current density $J_y\hat{\mathbf{y}}$ (not shown) are perpendicular to the magnetic, as they should be; they "wrap" around the magnetic field and "closes at $y = \pm$ infinity" since the model is 2D. The Alfvén wave corresponds to the right three panels, and its parallel propagation characteristics mentioned previously are apparent. The electric field is only in the x direction and drives the J_x component (not shown), which then closes through the field aligned current J_z. The magnetic field is purely in the y direction, and as expected is perpendicular to the electric field. On the ground, only the B_x component of the magnetic field survives. To get a feeling for the amplitude a magnetometer would measure on the ground ($z = 0$),

[10]If the magnetic field components for the MS wave were instead shown as a vector field superimposed on the field magnitude, the radiation pattern would look similar to the Figure 3.2c. The main difference, however, is that the MS waves in Figure 3.9 are ducted within the Alfvénic duct due to the plasma inhomogeneity, as was mentioned in Section 3.2.1.

we can take the discrete Fourier transform of the B_x time series by applying the well known Fast Fourier Transform (FFT) algorithm. The result of this is shown in Figure 3.10, where the FFT amplitude has been scaled to be in units of pT [11]. The

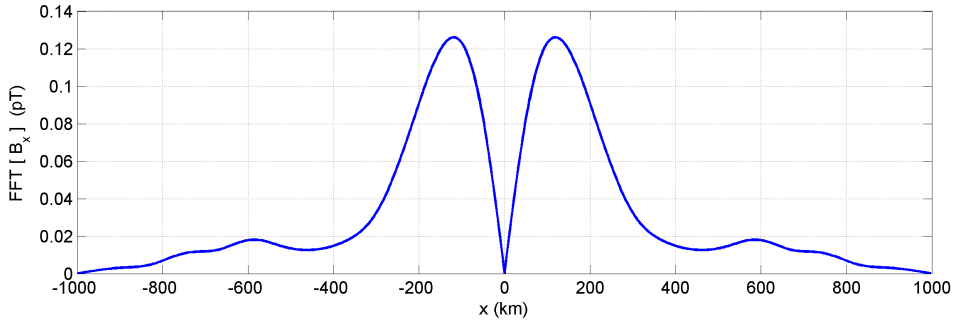

Figure 3.10: Discrete Fourier transform of B_x component, obtained via FFT, on the ground ($z = 0$) at each horizontal point in space (x). The FFT has been scaled in such a way that it is in units of pT.

FFT gives the discrete Fourier transform of the B_x component on the ground at each horizontal point in space, x. The magnetic field peaks near 120 km away from the center and has a secondary bump near 600 km, which is due to the MS wave interacting with the Hall region. Immediately below the source there is a null point since the vertical $\mathbf{B}_0$ in this simulation gives the system symmetry. Namely, the diamagnetic current in the heated region generates equal and opposite fields about the $x = 0$ line that cancel each other.

3.3.1 ICD Frequency Dependence

The proof of principle experiments that were reported by [37] (and mentioned in Section 3.2.2) showed that ground measurements of ICD signals slightly increased (or stayed close to flat) as the modulation frequency was increased from the low frequency end (about 0.5 Hz) to about 11-12 Hz. After this, a $1/f$ type of behavior

[11] To put it another way: with this scaling, the peak value of a spectrum, obtained from a pure sinusoidal waveform of amplitude unity, will itself be unity.

was observed, signifying that the modulation period was approaching the F region response time scale. It is therefore interesting to explore how the ground field amplitudes change as a function of the modulation frequency in the ICD model. However, it should be noted that the effects of the F region response time scale are not built into the model as it currently stands. While the F region response is an important feature that should be included in future work of the ICD model, it would be interesting to keep it fixed for all frequencies (as the current heating source (3.24) already does), and see how other factors affect the frequency dependence of ICD. Simulation results of the ICD frequency dependence will be discussed in the remainder of this chapter.

3.3.1.1 First Results: Source Size Dependence

To simulate the ICD frequency dependence, the ionospheric conditions are matched as closely as possible to the ones of the experiments in Reference [37]. To this end, we assume the spatial dependence of the plasma density is given by the Chapman profile (3.27), and set n_{max} such that the $f_0F_2 = 2$ MHz, as was approximately the case during most of the HF heating time. Moreover, the altitude of maximum plasma density (h_mF_2) was in the range 275 - 300 km, so z_{max} is set to 300 km. The altitude dependence of Γ_{en} and Γ_{in} are taken to be Equations (3.28) and (3.29), since they represent typical nighttime collision frequency profiles. Without an ISR at HAARP, accurate profiles for $n_0(z)$ are unfortunately not known. The same goes for $\Gamma_{en}(z)$ and $\Gamma_{in}(z)$, since accurate profiles for the neutral density and electron temperature are also unknown. We are using the plasma parameters and

profiles shown in Figure 3.8. Since HAARP operated at 2.8 MHz while the f_0F_2 was close to 2 MHz, the heating is expected to be collisional *underdense heating*. The altitude dependence of the temperature perturbation generated by the heating in this case should be quite extended in the vertical direction, peaking near the h_mF_2 [21]. Moreover, since the HAARP beam pattern expands outward roughly as a cone the horizontal width of the heated region will increase as a function of altitude. That being said, the temperature perturbation profile is again chosen to be Equation (3.24), but this time with parameters that more closely match the above mentioned properties: $z_{heat} = 275$ km (close to the h_mF_2), $D_z = 65$ km (extended in the vertical direction), and $D_x = (25, 34, 43)$ km at altitudes of $z = (150, 200, 250)$ km, with linear interpolation of D_x for all other altitudes. The D_x values are chosen in such a way that the angular width of the "cone" matches the HAARP angular beam width used by [57]. The conically shaped Gaussian temperature source is illustrated by Figure 3.11. Furthermore, the density gradient is neglected as in Equation (3.24),

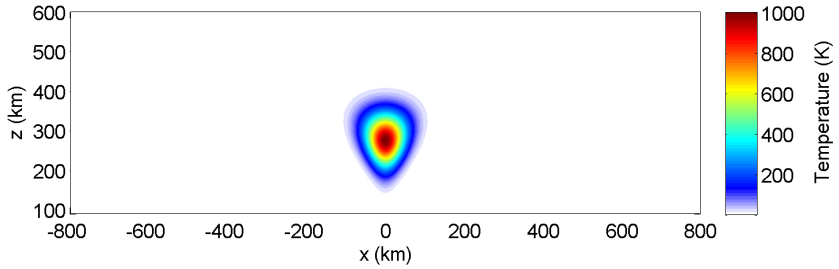

Figure 3.11: Temperature perturbation source with a vertically extended profile and conically shaped waist.

a vertical background field is assumed, D_t is set to 0.5 seconds, and T_{heat} is again set to 1000 K. For the given parameters, we do separate simulations for several modulation frequencies ($\omega = 2\pi f$) spanning some of the ULF and low ELF range: $f = 1$, 3, 5, 7, 9, 11, 15, and 20 Hz. Each frequency gives results similar to Figure

3.9. As an example, Figure 3.12a shows the magnetic field components at the end of the simulation ($t = 1$ sec.) for the case $f = 7$ Hz. The magnetic field components

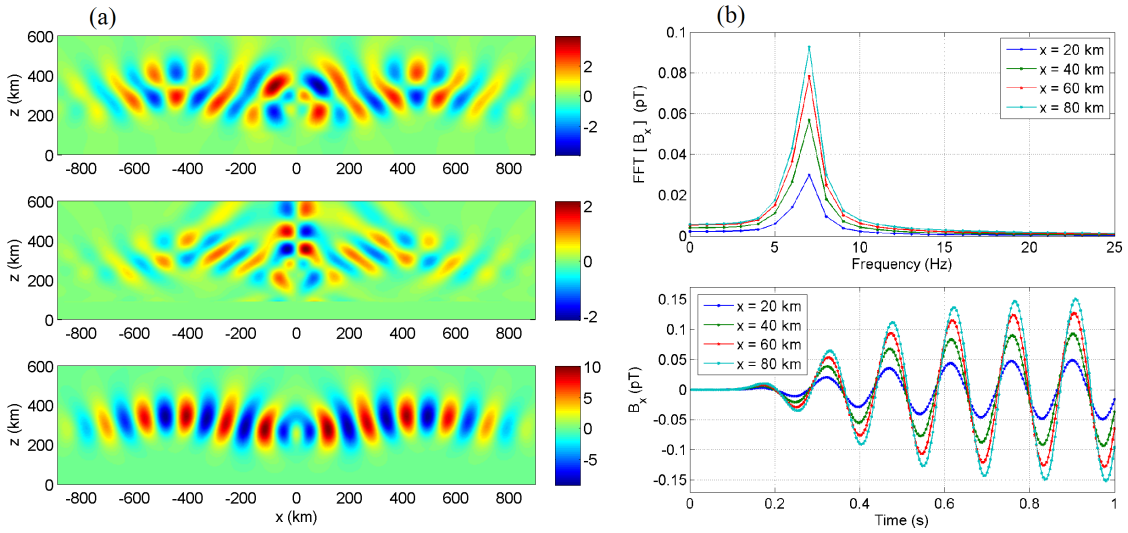

Figure 3.12: (a) Magnetic field components B_x (top), B_y (middle), and B_z (bottom) in pT at the end of a simulation ($t = 1$ second). (b) Frequency spectrum of B_x in pT at various observations points (top) obtained by taking the FFT of the B_x temporal waveform observed at those observation points (bottom).

B_x (top), B_y (middle), and B_z (bottom) and presented in pT and show a similar field distribution to before. On the ground, we can pick an observation point (say, to the right of $x = 0$) to see how the B_x component evolves in time, and subsequently apply the FFT to see its frequency spectrum. We can scale the FFT such that the spectral amplitude is in pT, as was done for Figure 3.10. Picking the observation points $x = 20, 40, 60, 80$ km, we obtain the results shown in Figure 3.12b. The top panel of the figure shows the frequency spectrum of B_x in pT at the observations points mentioned above, while the bottom panel shows the B_x waveform from which the spectrum was obtained. The peak in the frequency spectrum at 7 Hz is quite clear, and expected. Observation points to the left of $x = 0$ are not shown since they are by symmetry the same, and the $x = 0$ case is not shown since that is the null point. By doing this for each of the simulations corresponding to the remaining

frequencies and extracting the spectral amplitude for each case, we can compile them together to obtain a figure that illustrates the frequency dependence on the ground at the various observation points, as is revealed by Figure 3.13a. With the setup

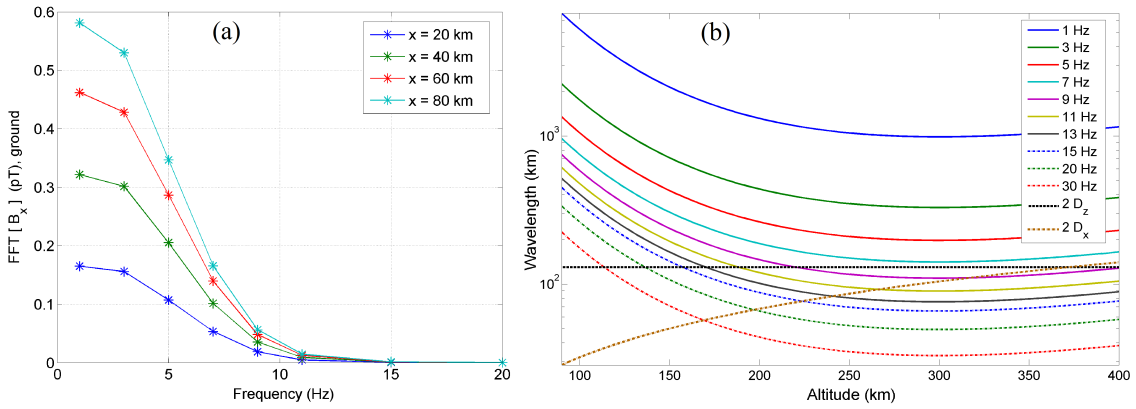

Figure 3.13: Spectral amplitude of the B_x component vs. modulation frequency as "measured" on the ground at various observation points ($x = 20, 40, 60, 80$ km). (b) Wavelength vs. altitude for various frequencies, where the vertical and horizontal source sizes have been superimposed for comparison.

described above, the ICD model predicts a steady drop-off of the field amplitudes measured on the ground as the frequency of the generated waves is increased. Note also that the amplitude of the higher frequencies, such as the 15 and 20 Hz cases, become essentially imperceptible compared to the lower frequencies.

The magnetic field amplitudes in Figure 3.13a are not in agreement with the experimental results; the amplitude is dropping above 5 Hz, rather than increasing or staying flat, as is suggested by experiments. Moreover, experiments observed signals for frequencies up to 60-70 Hz, while the model results imply frequencies above approximately 15 Hz would be unmeasurable. The behavior exhibited by the model, as it turns out, is essentially dictated by the relative sizes of the temperature source and the local wavelength of the generated waves. Namely, the diamagnetic current "antenna" that generates the MS waves is a poor radiator when the source size is

comparable with or bigger than the MS wavelength; (see Section 3.3.2 for details.) Since D_x and D_z are the half widths of the temperature source, we can estimate the horizontal and vertical sizes as $2D_x$ and $2D_z$ and plot them simultaneously with the altitude dependent Alfvén wavelength (λ), as is shown by Figure 3.13b. Roughly speaking, if the ratio $\lambda/(2D_x)$ is too small then the source is a poor radiator in the horizontal direction, while if the ratio $\lambda/(2D_z)$ is too small then the source is a poor radiator in the vertical direction. The figure reveals that near the peak of the source at 275 km, the wavelengths corresponding to frequencies 9 Hz and greater should be poor at radiating vertically. On the other hand, approximately 11 Hz and greater should be poor at radiating horizontally. A look back at Figure 3.13a corroborates these notions, since frequencies near or above 9 Hz have significantly smaller amplitudes than the lower frequencies.

It would be interesting to see how the results change when the source size is decreased in the vertical direction. Keeping everything in the simulations the same, but changing the vertical half-width to $D_z = 20$ km gives a ground amplitude response shown by Figure 3.14. It is quite clear that the field amplitudes for the

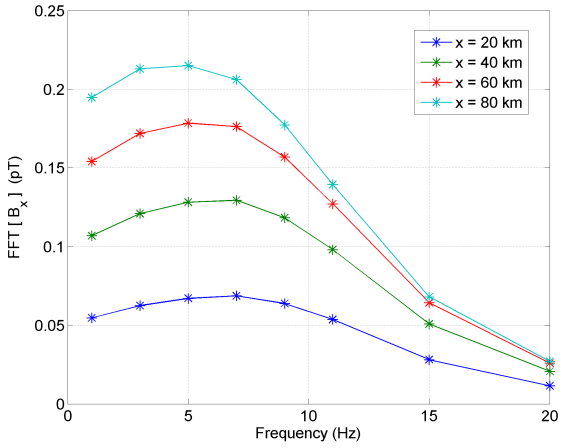

Figure 3.14: Same as Figure 3.13a, except the vertical width has been decreased to $D_z = 20$ km.

higher frequencies have gone up considerably. Interestingly, the amplitudes of the lower frequencies have gone down a little. This change is actually somewhat of a different effect than those discussed so far, and is more straightforward to understand. Since the lower frequencies essentially always have wavelengths larger than the source, they are not plagued by the same "source inefficiency" that affects the higher frequencies. Rather, as we shrunk the vertical source size from $D_z = 65$ to $D_z = 20$, we effectively reduced the "energy content" of the source, since the total energy content is proportional to the product of the pressure and the volume (actually, the area since the model is 2D). Therefore we would expect the amplitude to drop approximately by a factor of $20/65 \simeq 0.3$, which is consistent if you compared, say, the 1 Hz cases from Figures 3.13a and 3.14. While the end result is not quite what the experimental observations have suggested, the curves in Figure 3.14 do look more promising than those in Figure 3.13a.

3.3.1.2 Effects of ∇n_0 and Tilted Geomagnetic Field

The last section gave some preliminary result regarding the frequency dependence of ICD generated field amplitudes, as seen by an observer on the ground. Moreover, the results established that field amplitudes are quite sensitive to the ratio of the wavelength and source size. It would be fruitful to see how ICD is affected by a few other parameters of the system. For example, how would the results look if ∇n_0 was not neglected as it was in the previous simulations (Equation 3.24); or how would the asymmetry of a tilted background field change the ground field structure. To have a better understanding for how the density gradient affects the results, we

move the source from the previous 275 km to $z_{heat} = 200$ km where the gradient is larger (see Figure 3.8). For simplicity we keep the same source size used in Figure 3.14; namely, the conically shaped Gaussian with $D_x \approx 35$ km and $D_z = 20$ km. Performing a frequency sweep like usual and repeating the same procedure while also including the $\boldsymbol{\nabla} n_0$ term, we obtain the results depicted by Figure 3.15. Com-

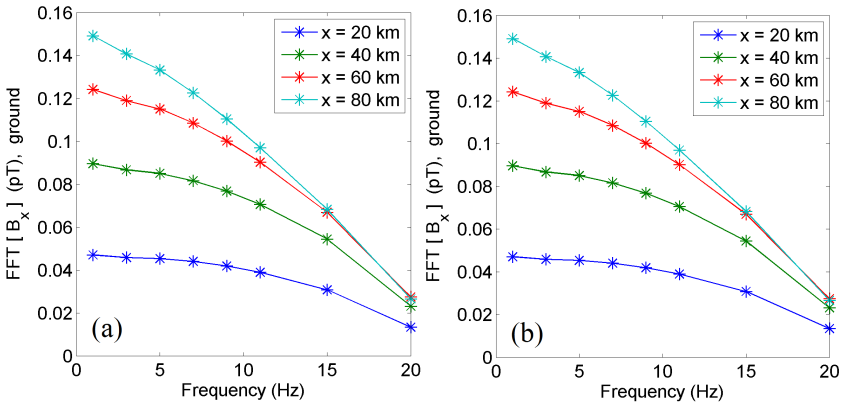

Figure 3.15: Simulation results of frequency sweep as before but with $z_{heat} = 200$, done for two cases: neglected density gradient (a) and included density gradient.

paring Figures 3.15a and 3.15b reveals that the $\boldsymbol{\nabla} n_0$ term had no effect on the field amplitudes on the ground. This result motivates a check of all contributions the source can make to wave generation, and in doing so the reason behind to above observation can be easily shown. First consider the source term in the ICD Equation 3.22b. Since this equation is an evolution equation for the electric field, the various vector components of the source terms will contribute to the corresponding components of the electric field. The source is given by

$$\left(\omega_{ci}\overline{\overline{\mathbf{R}}}_i - \frac{\partial}{\partial t}\right)\frac{\boldsymbol{\nabla} p_e}{e n_0},$$

which we can simplify since in the F region $\Gamma_{in} \simeq 0$, and thus write the tensor $\overline{\overline{\mathbf{R}}}_i(\cdots)$ as $\overline{\overline{\mathbf{R}}}_i(\cdots) \simeq (\cdots) \times \hat{\mathbf{b}}_0$, where $\hat{\mathbf{b}}_0$ is the unit vector in the direction of the background field. Factoring out the constant k_B/e and leaving it out for brevity, the above source becomes proportional to

$$\omega_{ci} \frac{\boldsymbol{\nabla}(n_0 T_e)}{n_0} \times \hat{\mathbf{b}}_0 - \frac{\partial}{\partial t} \frac{\boldsymbol{\nabla}(n_0 T_e)}{n_0}.$$

Since $\boldsymbol{\nabla}(n_0 T_e)/n_0 = \boldsymbol{\nabla} T_e + T_e \boldsymbol{\nabla} \ln(n_0)$, $\partial \ln(n_0)/\partial x = 0$, and $\partial / \partial y = 0$, the above becomes

$$\omega_{ci} \left[\hat{\mathbf{x}} \frac{\partial T_e}{\partial x} + \hat{\mathbf{z}} \left(\frac{\partial T_e}{\partial z} + T_e \frac{\partial \ln(n_0)}{\partial z} \right) \right] \times \hat{\mathbf{b}}_0 - \frac{\partial}{\partial t} \left[\hat{\mathbf{x}} \frac{\partial T_e}{\partial x} + \hat{\mathbf{z}} \left(\frac{\partial T_e}{\partial z} + T_e \frac{\partial \ln(n_0)}{\partial z} \right) \right]. \tag{3.30}$$

However, we used a vertical field for the simulation, i.e. $\hat{\mathbf{b}}_0 = \hat{\mathbf{z}}$, thus the above collapses to

$$-\hat{\mathbf{y}} \omega_{ci} \frac{\partial T_e}{\partial x} - \frac{\partial}{\partial t} \left[\hat{\mathbf{x}} \frac{\partial T_e}{\partial x} + \hat{\mathbf{z}} \left(\frac{\partial T_e}{\partial z} + T_e \frac{\partial \ln(n_0)}{\partial z} \right) \right]. \tag{3.31}$$

Note that now the density gradient appears only for the $\hat{\mathbf{z}}$ component of the electric field, which is a non-propagating component and does not contribute to wave dynamics. This observation accounts for the fact that the two cases shown in Figure 3.15 are identical.

While only the temperature gradient gives wave dynamics in this setup, note that if the background field was tilted away from the vertical in Equation (3.30), then we would get a nonzero contribution to wave dynamics from the density gradient. Letting $\hat{\mathbf{b}}_0 = \hat{\mathbf{x}} \sin\theta + \hat{\mathbf{z}} \cos\theta$, where θ is the tilt angle from the vertical, and focusing

only on the cross product term in Equation (3.30), we obtain

$$-\hat{\mathbf{y}}\omega_{ci}\left[\cos\theta\frac{\partial T_e}{\partial x}-\sin\theta\left(\frac{\partial T_e}{\partial z}+T_e\frac{\partial\ln(n_0)}{\partial z}\right)\right], \tag{3.32}$$

which explicitly shows the contribution from the density gradient.

Since it is now clear that only a tilted field will allow $\boldsymbol{\nabla}n_0$ to contribute, a set of simulation runs can be set up to do a frequency sweep. One setup is the already calculated case seen in Figure 3.15b. The second setup is the same ($\boldsymbol{\nabla}n_0$ included) but with the background field tilted at $\theta = 14°$. Figure 3.16 shows the results of these two cases side by side. Figure 3.16a shows the ground field amplitudes seen

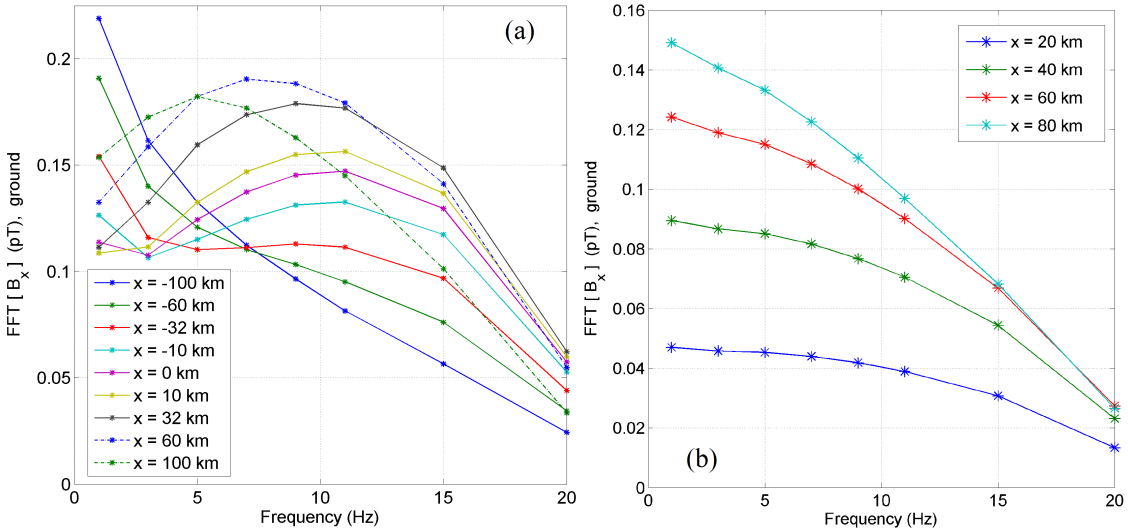

Figure 3.16: ICD simulations for the case of a tilted background field with the inclusion of $\boldsymbol{\nabla}n_0$ (a), and the previous result with a vertical field for comparison (b).

at various observation points as the modulation frequency is swept, for the case where the field has been tilted. Since the setup is now asymmetric, observation points to the left have been included as well. Figure 3.15b is reproduced here as Figure 3.16b to facilitate their comparison. The field amplitudes for the tilted case are noticeably stronger than before, though not by a very wide margin. In fact,

the observed amplitude differences are more due to the magnetic field orientation. A tilted field "directs" the MS waves toward to ground more than a vertical field, allowing more of the MS wave energy to reach the E region and leak towards the ground. The contribution of $\boldsymbol{\nabla} n_0$ is expected to be minimal for the following reason. By factoring out the temperature from (3.32), the magnitudes of the three terms

$$\left|\cos\theta \frac{\partial \ln(T_e)}{\partial x}\right|, \quad \left|\sin\theta \frac{\partial \ln(T_e)}{\partial z}\right|, \text{ and } \quad \left|\sin\theta \frac{\partial \ln(\mathrm{n_0})}{\partial \mathrm{z}}\right| \tag{3.33}$$

show that the relative contributions of each is determined by the length scales of the gradients of n_0 and T_e. The length scale of the ionosphere (and hence the plasma density) is on the order $\sim$ 100 km, while the length scale of the temperature gradient, say, in the x direction, is of the order $\sim$ 10 km (for the previous set of simulation). Moreover, $\sin\theta \simeq \sin(14^\circ) \simeq 0.25$, while $\cos(14^\circ) \simeq 1$, so the last term in (3.33) contributes about 2.5% compared to the first term. Therefore, a significant effect from the density gradient is not expected, even in the tilted case.

What remain to be considered are the time derivative terms in Equation (3.30) that have been left out from the analysis so far. However, most of the frequencies considered so far have been $\omega << \omega_{ci}$ ($\omega_{ci} \sim 40$ Hz); and $|\partial \,/\, \partial\, t| = \omega$, while the cross product terms in (3.30) are proportional to ω_{ci}. Therefore, the time derivative terms are not expected to contribute significantly in the presented simulations. That being said, they should have an increasingly important affect as the frequency is increased toward and above ω_{ci}.

Interestingly, Figure (3.16) shows a rise in ground amplitude starting from the lower frequencies, reaches a peak near 11-12 Hz, and decreases subsequently.

This is a good sign, since it is quite reminiscent of the experimental observations from Reference [37] mentioned earlier. The main discrepancy so far, however, is the fact that a relatively small source size was used to achieve this result, while the experimental conditions dictate that the source should be elongated in the parallel direction due to underdense heating. Moreover, frequencies higher than 20 Hz that have been experimentally observed, such as 50 Hz, have an wavelength of about 20 km in the heated region, thus requiring a source with $D_x < 10$ km in order to radiate efficiently; but, based on the known HAARP beam width this is not the case, and so the discrepancy remains unresolved. The following section addresses a possible resolution to this dilemma.

3.3.2 Wave Generation Efficiency of ICD Source

Section 3.3.1.1 claimed, and substantiated by numerical simulation results, that the source size-to-wavelength ratio in the ICD model has a significant influence on the field amplitudes. Since the simulation results include a host of effects, the reason behind the specific effect being mentioned is not transparent. The goal here is to use a simplified analytic model of wave generation to show the particular mathematical dependence that the wave amplitude has as a function of the source size-to-wavelength ratio. To this end, several approximations to the ICD model will be made to facilitate obtaining an analytic solution. Namely, since the frequencies considered in this work have been below the ion gyrofrequency, the MHD limit of the ICD model is assumed and yields Equations 3.26. We also take $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$, the ionosphere to be uniform ($n_0, B_0, v_A = \text{constant}$), and as usual the system is taken to

be 2D (i.e. uniform in the y-direction: $\frac{\partial}{\partial y} = 0$). Since the source is proportional to $\boldsymbol{\nabla} T_e \times \hat{\mathbf{z}} = -\hat{\mathbf{y}}\, \partial\, T_e / \partial\, x$, only the y-components of Equations 3.26a,b (corresponding to MS waves) are necessary. In this simplified setup, SA waves will not be generated since the homogeneity assumptions destroys the coupling between the two modes. However, this is of little concern since the MS electric field drives the SA waves, and thus knowledge of MS generation efficiency makes a statement about SA generation as well. Noting that $\hat{\mathbf{y}} \cdot [(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{A}) \times \hat{\mathbf{z}}] \times \hat{\mathbf{z}} = \nabla^2 A_y$, the y-components of 3.26 yield a nonhomogeneous *wave equation* for A_y:

$$\frac{\partial^2 A_y}{\partial t^2} = v_A^2 \nabla^2 A_y + \frac{k_B \omega_{ci}}{e} \frac{\partial T_e}{\partial x}; \quad T_e(x, z, t) = T_{heat} \sin(\omega t) \exp\left[-\frac{x^2}{D_x^2} - \frac{z^2}{D_z^2} \right],$$

while $A_x = A_z = 0$ since they have no driving source; the temperature source is taken to be harmonic in time and Gaussian distributed in space, as usual. The last major simplification is to take the source to be extended in the direction of $\mathbf{B}_0$ so that it primarily radiates in the perpendicular direction ($\hat{\mathbf{x}}$). Thus, by taking our observation point to be somewhere along $z = 0$, we can disregard the field amplitude dependence on z and simplify the above to a one dimensional wave equation given by

$$\frac{1}{v_A^2} \frac{\partial^2 A_y}{\partial t^2} = \frac{\partial^2 A_y}{\partial x^2} + \frac{k_B \omega_{ci}}{e v_A^2} \frac{\partial T_e(x, t)}{\partial x}; \quad T_e(x, t) = T_{heat} e^{-x^2/D^2} \sin(\omega t), \tag{3.34}$$

where the subscript on D has now been dropped for brevity. The source has been set to initially start at zero ($p_{ext}(x, 0) = 0$), though this is not strictly necessary since what we are really interested in is the *late-time* solution of Equation (3.34), when all

transient behavior induced by the initial conditions has cleared away. Disregarding the transient behavior simplifies the calculation further, since it allows us to factor out a harmonic time dependence for A_y. But first, to make calculations cleaner, the harmonic time dependence of the source is rewritten as

$$T_e \rightarrow -T_{heat} e^{-x^2/D^2} e^{-i\omega t}. \tag{3.35}$$

Since $\Im\left(-\exp(-i\omega t)\right) = \sin(\omega t)$, where $\Im(.)$ denotes the imaginary part, all we have to do is solve (3.34) with the source (3.35) and then take the imaginary part of the resulting A_y at the end of the calculation to obtain the true solution. With that established, we factor out a harmonic dependence from A_y, i.e. $A_y(x,t) \rightarrow A_y(x)\exp(-i\omega t)$, with the understanding that the true A_y is given by $\Im(A_y(x,t))$. With this, the wave equation (3.34) simplifies to a time-independent equation:

$$\frac{d^2 A_y(x)}{dx^2} + k^2 A_y(x) = C\frac{d}{dx} e^{-x^2/D^2}, \tag{3.36}$$

where we have defined the wavenumber $k \equiv \omega/v_A$, and the constant $C \equiv \omega_{ci} k_B T_{heat}/(e v_A^2)$.

Equation 3.36 is a harmonic oscillator equation with a driving source, and can be solve by the Green's function method. The Green's function, $G(x,x')$, for Equation (3.36) satisfies the the harmonic oscillator equation with a delta function source,

$$\frac{d^2 G(x,x')}{dx^2} + k^2 G(x,x') = \delta(x-x'), \tag{3.37}$$

and the solution is constructed by taking an integral over the product of $G(x,x')$

and the source, with respect to x':

$$A_y(x) = \int_{-\infty}^{\infty} G(x,x') \left(C \frac{d}{dx'} e^{-x'^2/D^2} \right) dx'. \tag{3.38}$$

There are two convenient Green's functions for Equation (3.37), $G_+(x,x')$ and $G_-(x,x')$, given by

$$G_\pm(x,x') = \mp \frac{i}{2k} e^{\pm ik|x-x'|}, \tag{3.39}$$

where the $(+)$ case is for a rightward going wave and the $(-)$ for a leftward going wave[12]. Choosing between $G_+(x,x')$ and $G_-(x,x')$ is essentially determined by the problem at hand and the preference of the user. We can arbitrarily *choose* to "observe" the wave to the right of the origin, and we would thus expect waves to originate at the origin and propagate to the right. Given this chosen observation point we must insist on using $G_+(x,x')$ from (3.39), at which point Equation (3.38) becomes

$$A_y(x) = -\frac{iC}{2k} \int_{-\infty}^{\infty} e^{ik|x-x'|} \left(\frac{d}{dx'} e^{-x'^2/D^2} \right) dx'. \tag{3.40}$$

The final step is to take the *far field* limit of Equation (3.40), since what concerns us is how much of the wave "makes it out" of the source (the near field) and travels into the so-called *radiation zone* (i.e. the far field). Mathematically, we take the limit $x >> D, \lambda$, i.e. we take the observation point, x, to be much larger than both the source size D and the wavelength of the wave, λ. Note that x appears

[12]Green's functions are not unique; a new one can be obtained by adding arbitrary homogeneous solutions of the wave equation. Incidentally, we can actually construct another Green's by adding $G_+(x,x')$ and $G_-(x,x')$ like this: $aG_+(x,x') + bG_-(x,x')$, for any $a+b=1$. Taking (say) $a=b=1/2$, the new Green's function, $\sin(k\,|x-x'|)/(2k)$, is *perfectly legitimate* but includes *both* right and left going waves, which is not as convenient for our purposes here.

inside an absolute value; but for *large values* of x, i.e. $x >> D$, we can approximate $|x - x'| \simeq (x - x')$ since when it flips the sign, the integrand is essentially *zero* there anyway, courtesy of the source (and its derivative) being localized to $|x'| \lesssim D$. This allows us to approximate the above integral as

$$\simeq \int_{-\infty}^{\infty} e^{ik(x-x')} \left(\frac{d}{dx'} e^{-x'^2/D^2} \right) dx' = ik \int_{-\infty}^{\infty} e^{ik(x-x')} e^{-x'^2/D^2} dx',$$

where the last step used integration by parts, and the boundary terms vanished since the source is zero at $\pm\infty$. Completing the square in the exponent allows us to easily integrate the above, and we obtain the following complex A_y:

$$A_y(x) = \frac{C}{2} D \sqrt{\pi} e^{-\frac{1}{4}k^2 D^2} e^{ikx}.$$

Putting the time dependence $\exp(-i\omega t)$ back in and taking the imaginary part, we finally obtain the far field of the real A_y:

$$A_y(x,t) = \frac{1}{2} C \sqrt{\pi} D e^{-\frac{1}{4}k^2 D^2} \sin(i(kx - \omega t)). \tag{3.41}$$

The dependence of the far field amplitude on the wavelength (λ) can be easily read off from Equation (3.41), and is given by

$$\text{far field } \propto \ \mathrm{e}^{-\frac{1}{4}\mathrm{k}^2\mathrm{D}^2} = \mathrm{e}^{-\pi^2/\bar{\lambda}^2}, \tag{3.42}$$

where $\bar{\lambda} \equiv \lambda/D$ is wavelength-to-source-size ratio. For easier interpretation of the far field dependence, Equation (3.42) is plotted in Figure 3.17a as a function of λ/D.

Similarly, Figure 3.17b shows a plot of (3.42) with respect to D/λ, but superimposed

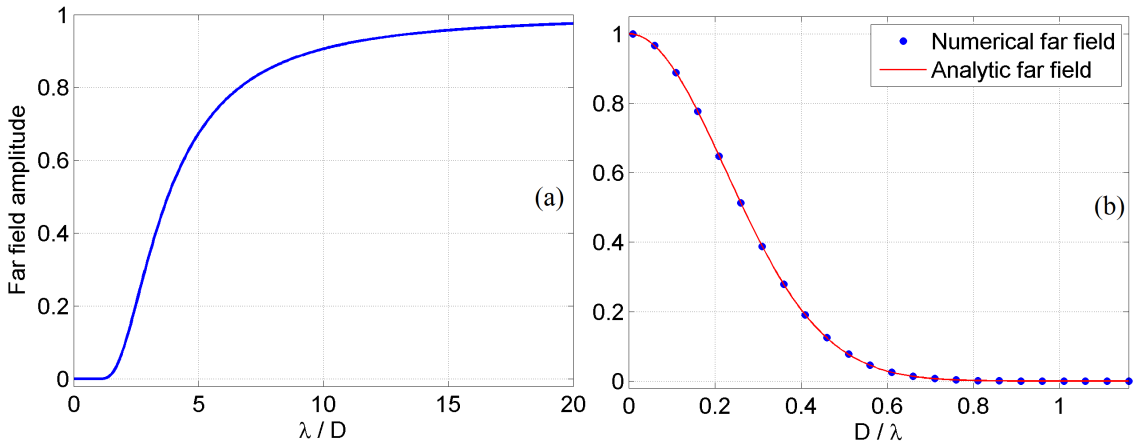

Figure 3.17: (a) Plot of Equation (3.42) showing the sudden drop of radiated amplitude for $\lambda/D \lesssim 1$. (b) A comparison of Equation (3.42) vs. D/λ (solid line) with numerically computed values (blue dots), showing that the analytic expression is quite accurate.

with values obtained from numerical solutions of Equation (3.34) (as seen in the far field); this illustrates that despite the several approximations used to get to this stage, the final result is quite accurate.

Note that for $\bar{\lambda} = \lambda/D \simeq 1$, we obtain $\simeq 5 \times 10^{-5}$ for the far field amplitude (relative to its maximum value). Even for the case when the wavelength is approximately the full size of the source, $\lambda \simeq 2D$, we obtain $\simeq 0.1$, one tenth of the maximum efficiency. It takes $\lambda \simeq 3D$ to obtain about one third of the maximum efficiency, after which the efficiency quickly rises for larger values of λ. This explains, at least qualitatively, why the waves being generated from the ICD model would experience such a drastic drop of their amplitude for higher frequencies. The initial ICD results shown in Figure 3.13a in fact look strikingly similar to Figure 3.17b. Despite the above analysis being somewhat artificial in its simplicity, the derived wave generation efficiency property is quite ubiquitous and will be qualitatively similar in much more general contexts, such as the full ICD model.

In the ICD model, T_e is assumed to be smooth with no variations within the Gaussian envelope. Despite temperature data being unavailable, it is reasonable to expect some spatial variations in the temperature on a scale that is smaller than the size of the source. In fact, the ionospheric plasma quite often becomes stratified along the background field from both HF heating and natural causes. A very crude way to see the potential effect of a stratified source is to perform the same analysis as above, but with a stratified external temperature source:

$$T_e(x,t) = T_{heat} e^{-x^2/D^2} \cos^2\left(\frac{\pi x}{L}\right) \sin(\omega t), \tag{3.43}$$

where L is the distance between the equally spaced stratified contributions. Carrying out the same procedure as before yields the result for the far field:

$$\text{far field} \;\propto\; \frac{\left[\mathrm{e}^{-\pi^2/\bar{\lambda}^2} + \frac{1}{2}\left(\mathrm{e}^{-\pi^2\left(\bar{\lambda}^2/\bar{\mathrm{L}}^2+1\right)^2/\bar{\lambda}^2} + \mathrm{e}^{-\pi^2\left(\bar{\lambda}^2/\bar{\mathrm{L}}^2-1\right)^2/\bar{\lambda}^2}\right)\right]}{1+\mathrm{e}^{-\pi^2/\bar{\mathrm{L}}^2}}. \tag{3.44}$$

The expression in the denominator is a normalization factor that keeps the area under $T_e(x)$ constant[13] as the value of L is changed (D is assumed fixed). This essentially ensures that the "energy content" of the source is the same for all values of L. Note that in the limit $L \to \infty$, the source approaches to the pure Gaussian case, and we exactly recover the far field (3.42) from before. A plot of Equation 3.44 as a function of $\bar{\lambda}$ for various values of L is shown in Figure 3.18. The behavior of the far field amplitude is quite interesting in this case since there is a drastic rise in wave generation efficiency for the smaller wavelengths. The $L \to \infty$ curve shown in

[13]It can be easily shown that $\int_{-\infty}^{\infty} e^{-x^2/D^2} \cos^2(\pi x/L) = \frac{1}{2}\sqrt{\pi} D(1 + e^{-\pi^2 D^2/L^2}) \propto 1 + e^{-\pi^2/\bar{L}^2}$.

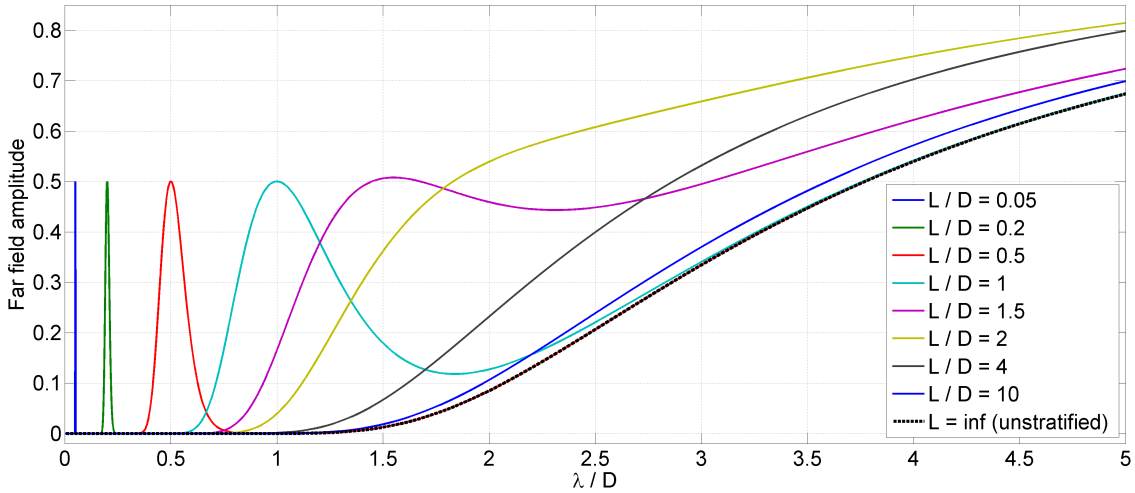

Figure 3.18: Far field amplitude dependence on λ/D for the case of a stratified external temperature source in Equation (3.34).

dotted black is the unstratified case, identical to Figure 3.17a. The cases with small values of L "hug" the unstratified curve everywhere except for $L \approx \lambda$, where there is a sharp peak, while the remaining values of L exhibit an intermediate behavior. The crucial difference between the smooth and stratified cases is the following: in the former case, newly generated waves from the smooth Gaussian source start to propagate out, but wavelengths smaller than the source size do not have enough time to "leave the source" before the harmonic time dependence flips its sign and generates waves that nearly cancel the original; these "phase cancellations" within the source result in a poor generation efficiency. The structure of the stratified Gaussian, however, allows for waves with wavelengths on the order of the striation size (L) to escape the source region with a significant amplitude. Finer structure in the source yields better generation efficiency for higher frequencies than the source would otherwise allow. The temperature distribution during ICD experiments is expected to have some fine structure as well, and thus this could explain how higher frequencies with perceptible amplitudes have been generated and measured on the

ground.

Consideration of self-consistent ways to include fine structure and its effects into the ICD model will have to be left to future work, along with another important effect not considered here: the F region temporal response to underdense heating. Progressing the current understanding of both of these effects would be greatly aided by an ISR, which as of yet is not available at the HAARP facility. There is, however, one final set of frequency sweep simulations that can give a better understanding of the remainder of the effects, which is achieved by suppressing the source inefficiency effects. To do so, the source size is made smaller than in the previously presented results. Setting the source size to 10 km $\times$ 10 km allows wavelengths for frequencies up to 30 Hz to radiate efficiently. Setting the background field to be vertical, the heating altitude to z_{heat} = 275 km, and keeping the rest of the model parameters the same as before, a frequency sweep of the model generates the ground field amplitudes shown by Figure 3.19. This is effectively the "point source" behavior of

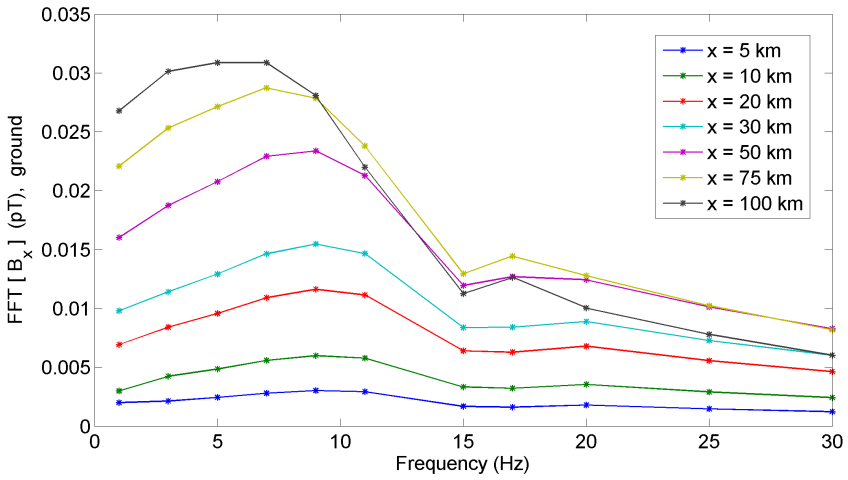

Figure 3.19: "Point source" behavior of the ICD model, obtained by taking a small heating source of size 10 km $\times$ 10 km; ground field amplitudes are shown as a function of modulation frequency for various observations points.

the ICD model as a function of the modulation frequency. Before interpreting the

results, a comment should be made with regard to the absolute amplitudes observed here. Namely, the highest amplitude seen in Figure 3.19 is about 0.03 pT - much lower than in the previous cases. Since the temperature perturbation amplitude has not been changed ($T_{heat} = 1000$ K) but the source size is reduced significantly, the energy content of the source has been reduced as well and the generated waves have smaller amplitudes as a result. In reality, if the HF energy was being pumped into a region as small as 10 km $\times$ 10 km then a more realistic temperature would be $T_{heat} = 5000$ K (as in Reference [39]). This would scale everything in Figure 3.19 by a factor of 5 and yield more familiar values for the ground amplitudes. That being said, the important features to take away are with regard to the relative amplitudes rather than the absolute ones. Note that the relative amplitudes are significantly altered compared to the initial ICD results, since all frequencies are now generated with roughly equal efficiency. The amplitude has a peak at the lower frequencies (5 - 10 Hz), along with a linear drop-off at higher frequencies ($\gtrsim$ 20 Hz). However, this drop-off is quite slow (and nearly flat in some cases), and would give reasonable amplitudes even for higher frequencies, such as 50 - 60 Hz.

3.4 Conclusions

The concept of the ICD method for generating ULF/ELF waves by modulated ionospheric heating of the F region was presented. Simulations of the mathematical model of ICD showed the generation of magnetosonic and shear Alfvén waves consistent with previous work. Sweeping some of the parameter space of the model revealed a sensitive relationship between the source size and wavelength of generated

waves. Namely, based on the Gaussian heating source used in the model, simulations predicted a severe drop in wave amplitude for wavelengths on the order of, or smaller than, the characteristic size of the source. The absence of this sensitivity in experimental observations motivated the analytic analysis of (simplified) ICD wave generation, which suggests that fine structure within the Gaussian envelope could explain the discrepancy. Moreover, a set of frequency sweep simulations showed that field amplitudes are not nearly as sensitive to other model parameters, such as the magnetic field orientation and plasma density. Several frequency sweep simulations revealed that the frequency dependence on the ground is qualitatively similar to the experimental observations, but only if a sufficiently small source size is assumed. These results, along with the simplified analytic analysis, suggest that a more accurate, first principles, description of the ICD source is necessary before consistency with the experimentally observed frequency dependence can be fully established.

Chapter 4

Whistler Wave Generation by Continuous Heating of the Upper Ionosphere

The generation of electromagnetic waves in the ELF/VLF frequency range by modulated HF heating of the ionospheric plasma has been the subject of many studies. Experiments and theory revealed that two completely different physical processes control their generation. The first process is electrojet current modulation, mentioned during the introduction of Section 3.2, and relies on modulated HF heating of the D/E region plasma electrons in the presence of electrojet currents. The resultant conductivity and associated electrojet current modulation generates a "virtual antenna" in the heated region (i.e. the PEJ antenna) that injects ELF/VLF waves in the earth-ionosphere wave guide, as well as whistler and shear Alfvén waves in the magnetosphere and the radiation belts. The major drawback of the PEJ antenna is the requirement of strong electrojet currents that render the process inapplicable in mid-latitude regions and with spotty reliability in the polar region. The second process, ICD, was the topic of Section 3.2.1, and relies on modulated heating of the electron pressure in the F region, without a need for electrojet currents. Due to the slow response time of the F region, the operation of the virtual ICD antenna

is limited to frequencies below about 60-70 Hz. Experiments indicate that in addition to ELF waves injected in the earth-ionosphere waveguide, ICD injects shear Alfvén waves in the magnetosphere along the Magnetic Zenith (MZ) of the heater and magnetosonic waves in the Alfvénic duct [39, 37, 14].

The objective here will be to present observations indicative of a third virtual antenna mechanism that injects broadband ($\Delta f/f \approx 0.1$ - 0.25) whistler waves into the magnetosphere. Whistler waves in the frequency range 7-10 kHz and 15-19 kHz, generated by F region CW HF ionospheric heating in the *absence* of electrojet currents, were detected by the DEMETER satellite overflying the HAARP transmitter during two ionospheric heating experiments. The whistler waves are in a frequency range corresponding to the F region Lower-Hybrid (LH) frequency and its harmonic, believed to be generated by mode conversion of LH waves that were parametrically excited by HF-pump-plasma interactions at the Upper Hybrid (UH) layer. We will discuss the basic physics and present a model that conjectures: (i) The whistler waves observed at the LH frequency are due to the interaction of the LH waves with meter-scale field aligned striations; (ii) The whistler waves at twice the LH frequency are due to the nonlinear interaction of two counter-propagating LH waves. The model is supported by numerical simulations and is shown to be in good agreement with the observations. The discussions of this chapter are based on the author's work in Reference [56].

4.1 Diagnostic Instruments

This chapter has the broadest range of instruments used for collecting the experimental observations presented in the next section, and as such invites a brief discussion. Much of the observations are gather by the same instruments as in the previous chapters. Namely, the ionosonde is used to obtain ionograms that give information about the state of the ionosphere and its characteristic parameters, while the nearby ground magnetometer is used to asses geomagnetic activity. The Kodiak radar (when available) is used to check if there is significant radar backscatter due to pump-induced field-aligned density striations, which is an indication of strong ionospheric heating. The full arsenal of instruments onboard the DEMETER satellite is available as well. The most important of these for the purposes of this chapter is the ICE instrument that measures one component of the electric field in the VLF range. The last diagnostic tool, which is getting its mention in this dissertation essentially for the first time, is measurements of *stimulated electromagnetic emission*. This is a very important and widely used diagnostic in ionospheric physics that is briefly discussed below.

4.1.1 Stimulated Electromagnetic Emission

As the O-mode HF pump wave enters the UH region of the ionosphere, some of its energy is converted into ElectroStatic (ES) waves due to the presence of seed plasma density irregularities. At the reflection altitude the pump wave excites Langmuir waves, while at the UH altitude it excites UH waves. Continuous pumping

keeps this process going until the amplitude of the ES waves increase to a critical value whereby nonlinear wave-wave interactions become important. At its core, the nature of a wave-wave interaction is to create "beat waves" that *resonantly drive* other plasma wave modes at this beat frequency. These nonlinear interactions, which generate various additional frequencies different from the pump frequency, are called *parametric instabilities*.

A heuristic picture of parametric instabilities can be laid out as follows. Consider three coupled wave modes represented byA_0 (pump),A_1, A_2 with frequencies f_0, f_1, f_2. Now, assume the coupling term between A_2 and A_1 is proportional to the product of A_0 and A_1. The product A_0 and A_1 will result in a beat wave with frequency $f_0 - f_1$, which will drive A_2 at that frequency. If A_2 has an allowed range of frequencies that happens to include $f_0 - f_1$ then A_2 is resonantly driven and will rapidly grow. Amazingly, the parametric instability "picks out" the frequency matching condition $f_0 - f_1 = f_2$ because of this resonance. Writing this differently we have$f_1 = f_0 - f_2$, which is to say that A_1 has a frequency that is *downshifted* from the pump by f_2. To put it another way, the pump wave "decays" into two daughter waves in such as way that

$$\omega_0 = \omega_1 + \omega_2,$$

where the (more conventional) corresponding angular frequencies have been used instead. This frequency matching is effectively a statement of conservation of energy of "photons", and an analogous statement for conservation of momentum can be

made as well:

$$\mathbf{k}_0 = \mathbf{k}_1 + \mathbf{k}_2,$$

where $\mathbf{k}_0$, $\mathbf{k}_1$, and $\mathbf{k}_2$ are the corresponding wavevectors.

For the case of the HF pump wave interacting with the ionospheric plasma, many such parametric instabilities are possible due to the zoo of wave modes supported by the F region plasma. The most important parametric instability for the purposes of this chapter involve UH and LH waves. As the HF pump wave mode-converts to large amplitude ES UH waves (of the same frequency) near the UH altitude, they act as an ES pump that decays into another (daughter) UH wave and a LH wave, with matching conditions identical to the above:

$$\omega_{uh_0} = \omega_{uh_1} + \omega_{lh}; \quad \mathbf{k}_{uh_0} = \mathbf{k}_{uh_1} + \mathbf{k}_{lh}.$$

The daughter UH wave, which has a frequency downshifted from the pump by the LH frequency (i.e. $\omega_{uh_1} = \omega_{uh_0} - \omega_{lh}$) can mode-convert back to HF EM radiation, which leaves the heated region and propagates towards the ground. Moreover, this decay processes can continue and generate a cascade of daughter UH waves, each new one downshifted from the previous by the LH frequency, i.e. $\omega_{uh_1} = \omega_{uh_2} + \omega_{lh}$, where for simplicity the frequency of the LH wave is taken to be the same as before.

A similar process takes place for various other interactions involving different ES waves (e.g. Langmuir, ion acoustic, Bernstein modes). The collection of all the EM radiation leaving the heated region due to these processes is referred to as Stimulated Electromagnetic Emission (SEE). The SEE is detected on the ground

by receivers and Fourier analyzed to extract the various off-shifted frequency bands. Figure 4.1 shows an example of an SEE spectrum and labels the various common features. The main central peak corresponds to the backscattered pump wave, while

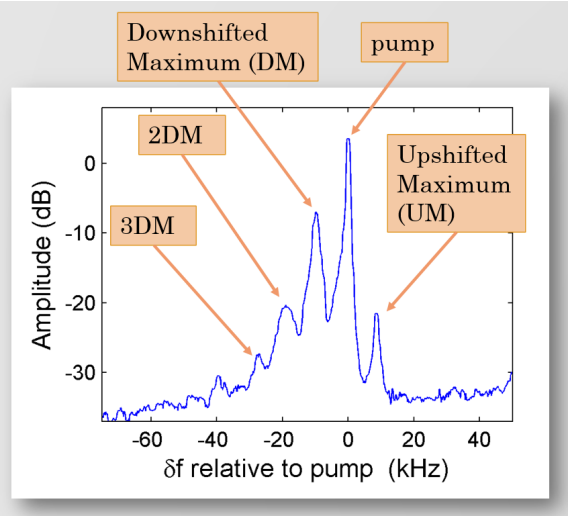

Figure 4.1: Power spectral density of SEE date showing typical SEE features (DM, 2DM, 3DM, and UM) observed on the ground for two different experiments.

the *Downshifted Maximum* (DM) is the feature downshifted relative to the pump frequency by the LH frequency (typically in the range $7 - 12$ kHz). The cascade of UH waves described above has also generated two more downshifted maxima (2DM, 3DM), and another feature called the Upshifted Maximum (UM) that is upshifted in frequency by approximately the same amount as the DM is downshifted.

4.2 Experimental Observations

We report below the results of two day-time HAARP heating experiments conducted during flyovers of the DEMETER satellite. In both experiments the HAARP heater operated at its maximum 3.6 MW power with O-mode, and the HF beam was directed along the HAARP MZ. The HF heating frequency was chosen to either match the critical frequency (f_0F_2) or coincide with the third electron

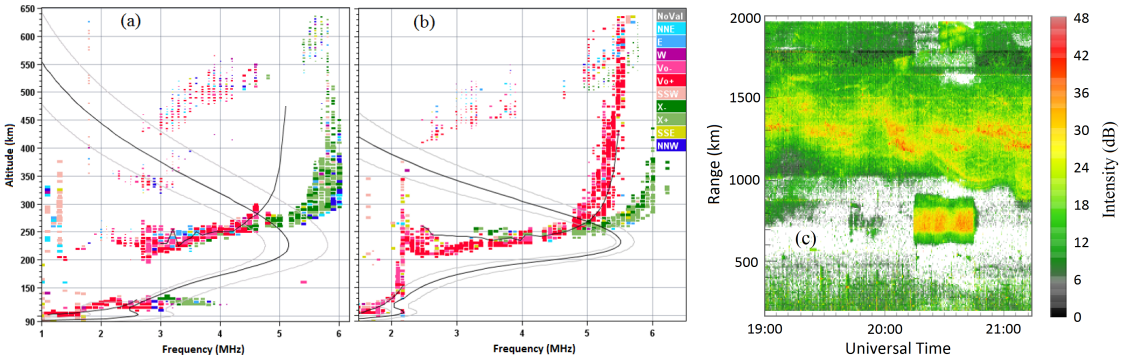

Figure 4.2: Ionograms during Exp. 1 (a) and during Exp. 2 (b) showing a smooth ionosphere, as well as Kodiak radar observations during Exp. 1 (c) showing strong reflections (given in dB) during the time of heating (20:15-20:45 UT).

gyro-harmonic (4.35 MHz). Key parameters of the experiments are given in Table 4.1, where the parameter labels used are identical to those in Table 2.1. Besides the HF frequency the only difference between the two experiments was that Exp. 1 (10/16/2009) used CW heating, while Exp. 2 (02/10/2010) used 0.7 Hz square pulse modulation. Both experiments were conducted during a smooth F region

Experiment label	Heating time (UT)	f_H (MHz) / h_r (km)	f_0F_2 (MHz) / h_mF_2 (km)	ΔR (km) / L_{EW} (km)	HF heating regime
1	10/16/2009 20:15-20:45	5.1 / 220	5.15 / 225	69 / 32	CW
2	02/10/2010 20:15-20:34	4.25 / 200	5.5 / 230	40 / 39	Modulated at 0.7 Hz

Table 4.1: Summary of key experimental information.

with weak-to-moderate D/E region absorption, as indicated by Figure 4.2a,b, and a steady ground magnetometer reading of about 10 nT (not shown). Figure 4.2c reveals Kodiak radar measurements during Exp. 1 and shows strong reflections during heating, indicating the build-up of strong plasma striations; the Kodiak radar was unavailable during Exp. 2.

The experimental results are shown in Figures 4.3 and 4.4. Figures 4.3a and 4.3b show spectrograms observed during DEMETER flyovers of HAARP at the

time of the experiments, with time measured relative to the closest approach of the MZ. The spectrograms were computed directly from the DEMETER ICE waveform

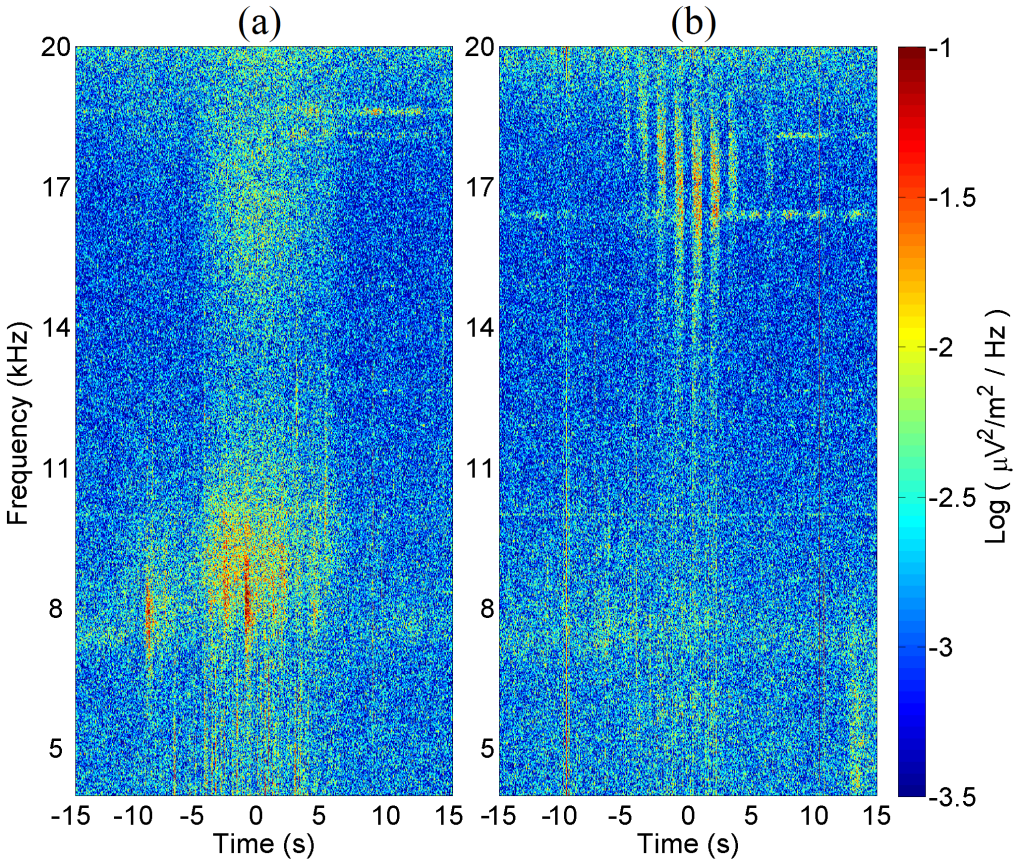

Figure 4.3: Spectrogram seen by DEMETER during Experiment 1 on 10/16/09 in which HAARP used CW heating (a), and during Experiment 2 on 02/10/10 (b) in which HAARP used 0.7 Hz square pulse modulated heating. In both cases time = 0 corresponds to the closest approach of DEMETER to the HAARP MZ.

data of one component of the electric field in the VLF range, by using a short-time Fourier transform with a window size of 4096 points and 50% overlap. In both figures strong whistler signals are observed in the vicinity of the MZ over approximately 10 s, corresponding to a distance of 75 km. Note that the regular temporal structure of the whistler waves in Figure 4.3b is attributable to the 0.7 sec on-off square pulse HF heating, while the irregular structure in Figure 4.3a can be attributed to CW heating in the presence of artificial ducts [59] (this is the topic of

Section 2.4.1). We should remark that the spectrograms contain features that are not related to our experiments: a faint band at $\sim$ 7 - 8 kHz stretching across the entire time domain, corresponding to naturally occurring LH oscillations; broadband and temporally narrow spectral features outside of the heated region due to short timescale transient processes, such as whistler waves created by lightning; strong but quite narrowband signals (e.g. at 10, 16.5 and 18.5 kHz) due to man-made transmissions.

Figures 4.4a,b show the power spectral density (PSD) estimates for the two experiments, measured near the closest approach to the HAARP MZ. For Exp. 1

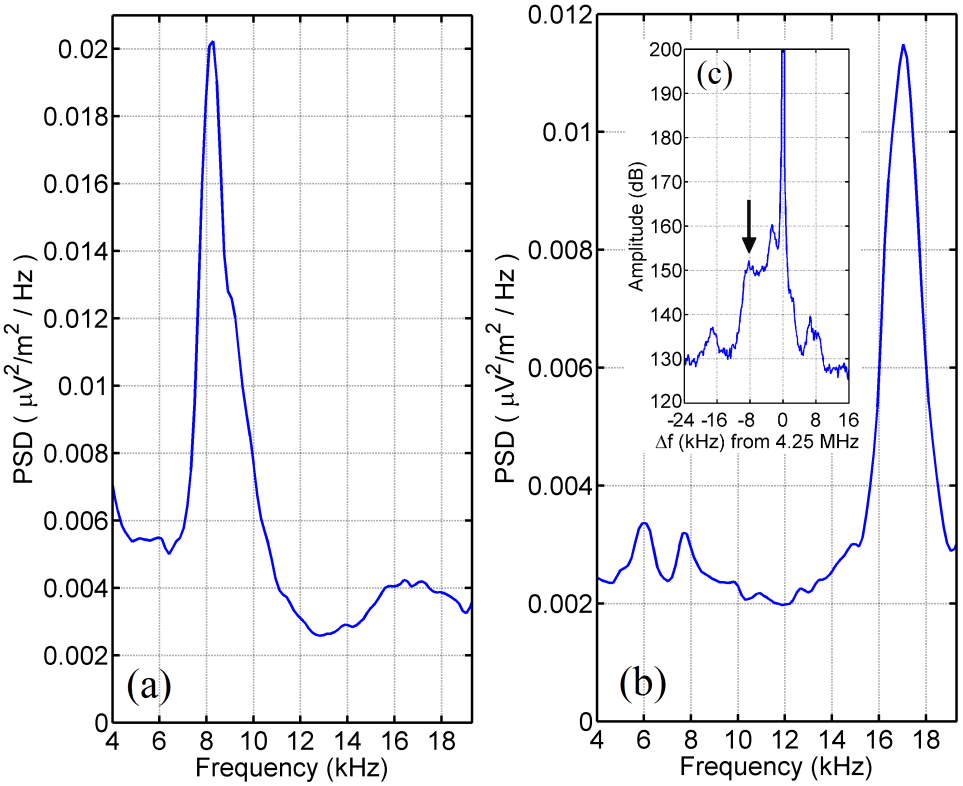

Figure 4.4: PSD of whistlers ($\mu V^2/m^2/Hz$) measured by DEMETER on 10/16/09 (a) and on 02/10/10 (b), calculated from 2 seconds of waveform at the closest approach, as well as the SEE spectrum observed on 02/10/10 (c); SEE amplitude is in dB, while Δf (kHz) is the frequency relative to the pump frequency (4.25 MHz)

(Figure 4.4a) the central frequencies are at 8.2 kHz and 16.5 kHz, corresponding to

the F region LH frequency and its harmonic, with Full Width at Half Maximum (FWHM) of approximately 2 kHz and 3 kHz. For Exp. 2 (Figure 4.4b) the central frequency is near 16.8 kHz, close to the LH harmonic, and has a 2 kHz FWHM. Note that the peak near 8 kHz in Figure 4.4b is of *natural* origin (as mentioned above) and *not* due to HAARP. However, the peak at 6 kHz is apparently due to the HAARP heating, as it is only present during on-times (see the very faint features near 6 kHz in the spectrogram 4.3b). This spectral feature, which is of unknown origin, is unrelated to the topics of discussion and as such will not be discussed henceforth. The PSDs in Figures 4.4a,b were obtained by applying Welch's method[1] to a 2 second interval near time = 0 in Figure 4.3a,b. In addition to the standard diagnostics, SEE measurements were available during Exp. 2. The inserted Figure 4.4c shows the SEE spectrum observed on the ground with the field amplitude in dB vs. Δf - the difference between the observed frequencies and the pump frequency. The large central peak is the backscattered pump wave, while the smaller peak (marked by the arrow) is the DM, downshifted from the pump by approximately 8.3 kHz. Note that the spectrum also shows a clear 2DM and an UM. In addition, there is the so-called "downshifted peak", downshifted by $2.5 - 3$ kHz. The relevance of SEE and the DM will become clear in the following section.

[1]This is a method of power spectral density estimation of a signal, computed as follows: (1) the signal data is split up into overlapping segments; (2) the data segments are windowed with an appropriate window function and their PSD is computed by applying the discrete Fourier transform; (3) the PSDs from all segments are averaged together to obtain the final result.

4.3 Discussion and Theoretical Considerations

The observed whistler waves are in a frequency range corresponding to the F-region LH frequency and its second harmonic. In this section we motivate that the observations are due to mode conversion of LH waves to whistler waves, where the LH waves were parametrically excited by the HF pump interacting with the plasma at the UH altitude. Moreover, in examining the observations presented in Figures 4.3 and 4.4 we find a major peculiarity: while in Exp. 1 whistler waves were measured at the LH frequency and its second harmonic, in Exp. 2 whistler waves appeared only at the harmonic of the LH frequency. To understand this puzzling absence and subsequent theoretical discussions, it pays to first review the HF pump wave-plasma interaction happening in the heated region.

4.3.1 Parametric Excitation of LH Waves at the UH Resonance

Consider what happens when an O-mode HF wave of frequency f_0 is transmitted along the MZ: the HF wave propagates to the reflection point where $f_0 = f_{pe}$ (the plasma frequency) and in addition encounter the UH resonance $f_0^2 = f_{pe}^2 + f_{ce}^2$ along the way, where f_{ce} is the electron gyrofrequency. At the UHR, the HF pump wave mode converts to UH waves due to natural or self-focusing driven [17] irregularities. UH waves are trapped in the irregularities and become amplified until parametric instabilities are triggered [18], thus exciting other wave modes. As explained in Section 4.1.1, the LH wave is manifested by the DM in the SEE spectrum and has been confirmed indirectly by numerous SEE observations. Over the years

several theoretical models have been proposed for DM generation, culminating in "cascade" models that naturally explain the multiple DM features (2DM, 3DM, $\cdots$) that are regularly observed during SEE experiments. For concreteness we assume a cascade model [52], summarized by the following step-by-step interaction, where the parenthesis indicate the frequency and perpendicular wavenumber of each quantity:

$$EM_{pump}(f_0, 0) + N(0, \pm k_N) \rightarrow UH(f_0, \pm k) \tag{4.1a}$$

$$UH(f_0, \pm k) \rightarrow UH_1(f_0 - f_l, \mp k_1) + LH(f_l, \pm k_l) \tag{4.1b}$$

$$UH_1(f_0 - f_l, \mp k_1) + N(0, \pm k_N) \rightarrow EM_{DM}(f_0 - f_l, 0) \tag{4.1c}$$

$$UH_1(f_0 - f_l, \mp k_1) \rightarrow UH_2(f_0 - 2f_l, \pm k_2) + LH(f_l, \mp k_l) \tag{4.1d}$$

$$UH_2(f_0 - 2f_l, \pm k_2) + N(0, \mp k_N) \rightarrow EM_{2DM}(f_0 - 2f_l, 0) \tag{4.1e}$$

$$\vdots$$

The above interactions can be described as follows: In the first process (4.1a) the dipole ($k_{em} \simeq 0$) EM pump wave (EM_{pump}) is mode converted on HF-pumped small scale striations of characteristic size $2\pi/k_N$ (N) and excites UH waves of the same frequency (and wavenumber $k \simeq k_N$), which get trapped inside the striations, leading to counter streaming waves ($\pm k$) and amplification of both the UH waves and striations. Once an UH wave reaches a threshold amplitude it parametrically decays by the 3-wave process (4.1b) into another UH wave (UH_1) and a LH wave. The UH_1 waves will be downshifted in frequency from the pump by the LH frequency, and by interacting with the striations they will mode convert back to EM waves by the mode conversion process (4.1c) and observed on the ground in the SEE spectrum

as a DM (EM_{DM}). As indicated by (4.1d) and (4.1e) this process can continue iteratively and generate a cascade of DMs (2DM, 3DM, $\cdots$) each new one further downshifted by the LH frequency from the previous. This is confirmed by the DM feature in Figure 4.4c being downshifted by about 8 kHz, and the 2DM being further downshifted by an additional 8 kHz; the presence of the DM during Exp. 2 is a proxy for parametrically excited LH waves.

4.3.2 Striation Development and the Missing LH Peak

Several theoretical and experimental studies [5, 15, 50] have shown that LH waves can be converted into whistler waves (W) (and vice versa) in the presence of meter-scale plasma density striations (D):

$$LH(f_l, \pm k_l) + D(0, \mp k_{str}) \rightarrow W(f_l, 0).$$

Following the start of HF heating, the development of SEE striations (N) and the DM (and hence LH waves) has been shown to be less than 20 ms [49], while the development of significant meter-scale sized striations take much longer, on the order of 5-10 s [23]. These times scales can be demonstrated by our recently conducted experiments at HAARP. Conditions were similar to those in Exp. 2 (daytime, quiet ionosphere, $f_H = 5.75$ MHz $\approx 4f_{ce}$, $h_r = 200$ km, pulsed MZ heating), and two complementary diagnostics were used: SEE measurements and GPS Slant Total Electron Content (STEC) data[2]; the latter method detects the effects due to artificial

[2]STEC is effectively the height integrated plasma density, which is obtained from overflying GPS satellites by measuring the phase delay between two GPS signals (of different frequency) following their propagation through the ionosphere. Density irregularities driven by ionospheric heating results in increased STEC, thus making it a useful diagnostic.

striations [34]. It was found that using a heating frequency that is 20 - 30 kHz below $4f_{ce}$ resulted in SEE with a well-developed DM and simultaneously an increase in STEC, corresponding to the formation of plasma density striations, as is shown by Figure 4.5. The STEC had a build-up time of about 5-10 s (Figure 4.5a), while the

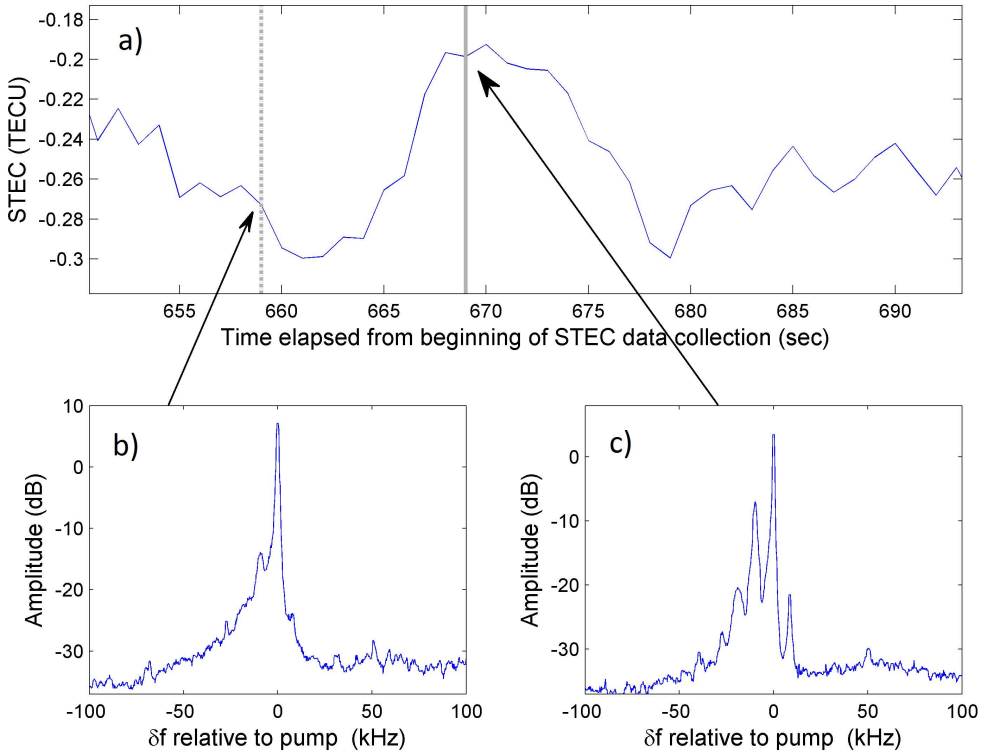

Figure 4.5: Measurements of STEC during an HF heating experiment (a); variation in STEC has a timescale of several seconds. Simultaneously, SEE was measured 20 ms after the start of heating (b), already showing signs of DM. SEE after 10 s of heating shows well developed DM, along with 2DM and 3DM (c).

buildup time of the DM was under 20 ms (Figure 4.5b), and became fully developed with multiple DM features (via the cascade model described above) after about 10 s (Figure 4.5c).

We are now in a position to explain the missing peak near the LH frequency in Figure 4.4b. Recall the main difference between the two experiments: Exp. 1 used CW heating, while Exp. 2 used pulsed heating with on/off times of 0.7 s.

Figure 4.5a,b illustrates that using short heating pulses with 0.7 s on and 0.7 s off, such as in Exp. 2, is long enough to generate LH waves, but too short to allow the development of artificial striations. Without a sufficient build-up of striations, the linear mode conversion mechanism would be too inefficient to be observed by DEMETER; this is consistent with the absence of VLF near the LH frequency in Figure 4.4b.

4.3.3 Whistler Waves at the LH Harmonic

The main peak in Figure 4.4a near the LH frequency is naturally explained by the fact that CW heating generates meter-scale striations necessary for LH-whistler conversion. Our attention now shifts to the LH second harmonic in Figure 4.4a,b. LH waves with a frequency near (or greater than) twice the LH frequency cannot exist, since this would break the LH existence requirement $\pi/2-\theta < \sqrt{m_e/m_i}$, where θ is the angle between the LH wave vector and the background geomagnetic field, $\mathbf{B}_0$. Thus the second harmonic must be generated from a different kind of interaction, presumably a nonlinear one due to the frequency change. We propose that the mechanism responsible for the LH harmonic is due to the nonlinear interaction of oppositely propagating LH waves, in analogy with counter-streaming Langmuir waves interacting to give EM waves with twice the Langmuir frequency [1]:

$$LH(f_l, +k_l) + LH(f_l, -k_l) \rightarrow W(2f_l, 0). \tag{4.2}$$

Such a mechanism does not directly rely on striations, but instead relies on the density fluctuations due to the large amplitude LH electric field. Now, consider two

oppositely traveling LH waves, one slightly oblique and one perpendicular to $\mathbf{B}_0$, with electric fields E_1 and E_2, respectively:

$$E_1 \sim \exp\left(ik_{l,\perp}y + ik_{l,\parallel}z - i\omega_l t\right); \quad E_2 \sim \exp\left(-ik_{l,\perp}y - i\omega_l t\right).$$

A nonlinear "beating" of these two waves gives

$$E_3 \sim E_1 E_2 \sim \exp\left(ik_{l,\parallel}z - i2\omega_l t\right), \tag{4.3}$$

which can mode convert to parallel whistlers if $\omega_w = 2\omega_l$ and $k_{w,\parallel} = k_{l,\parallel}$.

4.3.4 LH-Whistler Mode Conversion Model

Reference [15] studied LH-whistler mode conversion in the presence of plasma density striations by formulating the problem into two coupled equations, corresponding to the whistler and LH wave. The whistler wave was shown to be governed by the evolution equation for the whistler current

$$\frac{\partial \mathbf{j}_W}{\partial t} = -\frac{e\lambda_e^2}{m_e}(1-\lambda_e^2\nabla^2)^{-1}\boldsymbol{\nabla}\times\left[\boldsymbol{\nabla}\times\left((n_{str}+n_{LH})\mathbf{E}_{LH}+\mathbf{j}_W\times\mathbf{B}_0\right)\right], \tag{4.4}$$

while the LH current is governed by

$$\frac{\partial \mathbf{j}_{LH}}{\partial t} = \nabla^{-2}\left\{\frac{e}{m_e}\boldsymbol{\nabla}\times\left[\boldsymbol{\nabla}\times\left(n_{str}\mathbf{E}_W+\mathbf{j}_{LH}\times\mathbf{B}_0\right)\right]-\frac{e}{m_i}\boldsymbol{\nabla}\left[\boldsymbol{\nabla}\cdot\left(\mathbf{j}_{LH}\times\mathbf{B}_0\right)\right]\right\}. \tag{4.5}$$

(Note that the above "currents" are particle current densities, i.e. the product of number density and velocity.) The whistler and LH electric fields are given by

$$\mathbf{E}_W = -(\mathbf{j}_W \times \mathbf{B}_0)/n_0 \tag{4.6}$$

$$\mathbf{E}_{LH} = -\boldsymbol{\nabla}\left[\nabla^{-2}\left[\boldsymbol{\nabla}\cdot(\mathbf{j}_{LH}\times\mathbf{B}_0)\right]/n_0\right]. \tag{4.7}$$

For the above equations $\lambda_e = c/\omega_{pe}$ is the electron inertial length, $\mathbf{B}_0$ is the background field vector, n_0 is a constant background plasma density and n_{str} accounts for external density striations. Thus in the presence of density striations the LH electric field can drive whistler waves, and vice versa. Note that n_{LH}, which is the density fluctuation of the LH wave, was neglected in Reference [15] during linearization of the equations. We generalized the model by keeping this nonlinear coupling, for reasons that will become clear below. The density fluctuations are obtained from the continuity equation

$$\frac{\partial n_{LH}}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{j}_{LH} = 0. \tag{4.8}$$

In order to study the mode conversion of LH waves to whistlers, the author wrote a MATLAB code that numerically solves the above model consisting of Equations (4.4), (4.5), and (4.8); the spatial derivatives are approximated by pseudo-spectral derivatives[3], and the system is stepped forward in time with a standard fourth order Runge-Kutta method. As a reality check, the code was benchmarked by making simulation runs using input parameters and initial conditions identical to those presented in Reference [15] (with n_{LH} set to zero in the above model, of

[3]Pseudo-spectral derivatives are not only very accurate, but they are also quite fast, thanks to the FFT algorithm. Moreover, for the model at hand, they also have the advantage of turning the ∇^{-2} operators to division by $-k^2$ in Fourier k-space.

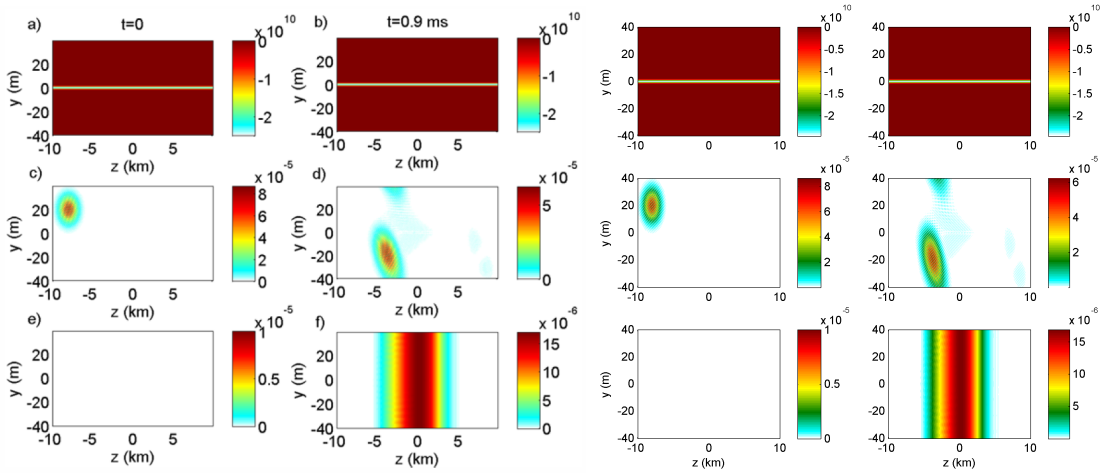

Figure 4.6: Simulation benchmarking results of LH-whistler conversion (right set of subfigures) being compared with those presented in Figure 1 of Reference [15] (left set of subfigures). Each is organized such that the left column is $t = 0$, while the right column is the end of the simulation at $t = 0.9$ ms; the rows (descending) are the density striation n_{str}, the LH wave amplitude, and the whistler wave amplitude.

course). Figure 4.6 shows simulation results of LH-whistler conversion, where the subfigures on the left are reproduced from Figure 1 of Reference [15], while the subfigures on the right are the results of the benchmark simulation run from the author's code, showing exact agreement. Each collection of subfigures is organized such that the left column is the initial setup ($t = 0$), while the right column is the simulation at $t = 0.9$ ms; the rows, from top to bottom, are the density striation n_{str} (which does not change), the LH wave amplitude, and the whistler wave amplitude. Notice that initially there are no whistler waves, but they are generated once the LH wave packet crosses the striation. Also note that the LH wave packet has deformed by the end of the simulation, since it is a dispersive wave.

Reference [15] also presented (in Figure 2) results of a set of parameter-sweep simulation runs, showing how the magnetic energy of the system (stored in the whistler waves) is affected by three parameters: (a) the angle (θ) between the LH wavevector ($\mathbf{k}$) and $\mathbf{B}_0$; (b) the perpendicular LH wavenumber, k; (c) the half-width

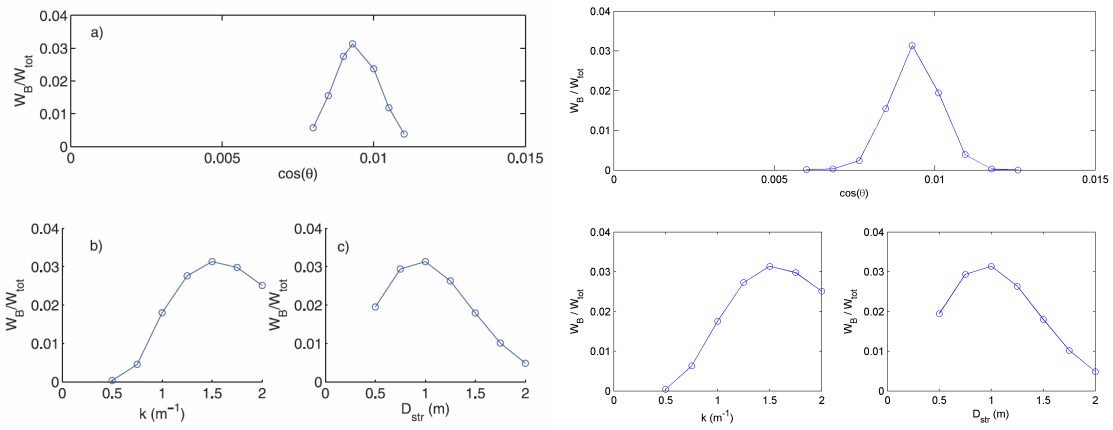

Figure 4.7: Parameter-sweep simulations for LH-whistler conversion (right set of subfigures) being benchmarked against those presented in Figure 2 of Reference [15] (left set of subfigures). The subfigures show how the magnetic energy of the system is affected by three parameters: (a) the angle (θ) between the LH wavevector ($\mathbf{k}$) and $\mathbf{B}_0$; (b) the perpendicular LH wavenumber, k; (c) the half-width of the striation, D_{str}.

of the striation, D_{str}. As a final benchmark for the code, Figure 4.7 shows a set of similar parameter-sweep simulations, alongside the results presented in Figure 2 of Reference [15]. Comparing them side-by-side shows good agreement, as one would expect.

Reference [15] showed that efficient resonant mode conversion from LH to whistler waves occurs when the striation half-width (D_{str}) and the perpendicular wavenumber of the LH wave ($k_{l,\perp}$) are related by

$$2D_{str} \approx \frac{\pi}{k_{l,\perp}}, \tag{4.9}$$

which is to say that the the striation full-width must be close to half the perpendicular LH wavelength. (This can be confirmed by either set of the bottom two subfigures in Figure 4.7.) During this resonant mode conversion process the LH and whistler waves have the same frequency: $\omega_l = \omega_w = \omega$. The wave vector components along the geomagnetic field are the same for the LH and whistler wave,

$k_{l,\parallel} = k_{w,\parallel} = k_\parallel$, but their perpendicular components $k_{l,\perp}$ and $k_{w,\perp}$ can be different. While the LH wave propagates almost perpendicular to the magnetic field, the whistler wave propagates primarily along the magnetic field but can be slightly oblique. For the discussions in this chapter, parallel whistler propagation ($k_{w,\perp} = 0$) is a good approximation and will be assumed below. With these conditions the LH and whistler dispersion relations give

$$\omega_l^2 = \omega^2 = \frac{\omega_{l,0}^2 k_{l,\perp}^2 + \omega_{ce}^2 k_\parallel^2}{k_{l,\perp}^2 + k_\parallel^2} \approx \frac{\omega_{l,0}^2 k_{l,\perp}^2 + \omega_{ce}^2 k_\parallel^2}{k_{l,\perp}^2}, \tag{4.10}$$

$$\omega_w^2 = \omega^2 = \lambda_e^4 k_\parallel^2 (k_{w,\perp}^2 + k_\parallel^2)\omega_{ce}^2 \approx \lambda_e^4 k_\parallel^4 \omega_{ce}^2, \tag{4.11}$$

where $\omega_{l,0} = 2\pi f_{l,0} = \sqrt{\omega_{ce}\omega_{ci}}$ is the LH oscillation frequency. Eliminating $k_\parallel^2$ in Equation (4.10) by using Equation (4.11), we then obtain

$$k_{l,\perp}^2 = \frac{\omega_{ce}\omega}{\lambda_e^2(\omega^2 - \omega_{l,0}^2)} = \frac{\sqrt{m_i/m_e}(f/f_{l,0})}{\lambda_e^2\left[(f/f_{l,0})^2 - 1\right]}. \tag{4.12}$$

Since the ionosphere has more than one ion species, m_i should be interpreted as an effective ion mass. By using the 2007 International Reference Ionosphere (IRI), we can estimate the local plasma parameters necessary for finding $k_{l,\perp}$ from Equation (4.12). We take the electron gyrofrequency to be $f_{ce} = 1.45$ MHz near altitudes of 200-220 km at HAARP [30]. For typical ionospheric conditions, such as in Exp. 1 and 2, the IRI model gives $f_{l,0} \approx 7.5$ kHz, $\lambda_e \approx 9$ m. If we take $f = 8.2$ kHz, corresponding to the main peak in Figure 4.4a, then Equation (4.12) gives an approximate range of $k_{l,\perp} \approx$ 3-4 m^{-1} for the relevant altitudes. The corresponding resonant striation width can be found from Equation (4.9) to be $D_{str} \approx 1$ m, which

is the characteristic size of small scale striations known to exist during continuous HF heating [11].

4.4 Simulation Results and Comparison with Observations

Consider the model (4.4) - (4.8) with input parameters similar to the above estimates. We take $n_0 = 3.1 \times 10^5$ cm^{-3} and assume a Gaussian depletion profile for n_{str} with a full-width $D_{str} = 0.8$ m and depletion amplitude that is 1.25% of n_0. Moreover, the initial conditions are set to be three LH wave packets directly on top of the striation, all with $k_{l,\perp} = 4$ m^{-1}. For the first wave packet, the angle (of the wave fronts) relative to $\mathbf{B}_0$ is set to resonantly generate whistler waves at the LH frequency by (linearly) interacting with the striation [15]. For the remaining two wave packets: one is set to be exactly perpendicular ($\theta = \pi/2$), while the angle of the other is (analogously) set to resonantly generate whistler waves at the LH harmonic by nonlinearly interacting with the perpendicular LH wave packet. Running a simulation with this setup generates mode-converted whistler waves with frequencies that correspond to the LH frequency and its harmonic, as shown in Figure 4.8. Figure 4.8a reveals the magnitude of the LH electric field vector, while Figure 4.8b shows the magnitude of the whistler magnetic field vector. The spectrum of the whistler magnetic field, as seen from a stationary observation point at the top of the simulation domain ($z = 9$ km), is plotted in Figure 4.8c and shows good agreement with the experimentally observed PSD in Figure 4.4a. The vertical lines indicate the LH frequency and its harmonic, and as expected they are close to the peaks of the whistler spectrum. If the striation amplitude were to be set to zero

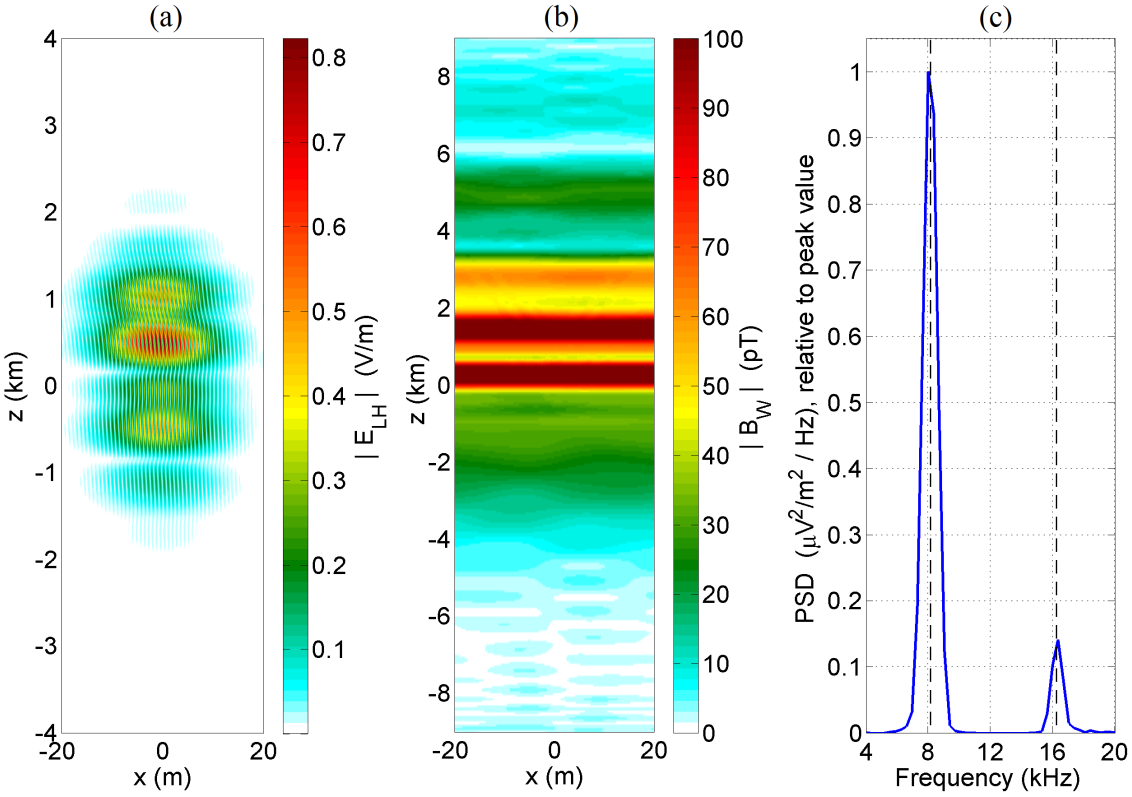

Figure 4.8: Simulation results showing generation of mode-converted whistler waves with frequencies that correspond to the LH frequency and its second harmonic: (a) The magnitude of the LH electric field in V/m, (b) magnitude of the whistler magnetic field vector in pT, and (c) PSD of the y-component of the whistler electric field, normalized to the peak value; the vertical lines in (c) represent the LH frequency and its second harmonic.

in the simulation, then only whistler generation at the LH harmonic would result, thus confirming the whistler spectrum observed Figure 4.4b.

4.5 Conclusions

This chapter described two HAARP/DEMETER experiments in which VLF waves of artificial origin were detected by the DEMETER satellite while overflying the HF-heated region of the ionosphere. The observations were shown to be consistent with parametrically excited LH waves being mode converted to whistler waves during HF heating. The VLF near the LH frequency observed during Exp. 1, which used CW heating, was shown to be due to resonant mode conversion to

whistler waves in the presence of artificially pumped meter-scale striations. The VLF near the LH harmonic observed during both Exp. 1 and Exp. 2 was shown to be generated by a different mechanism: the nonlinear 3-wave interaction of two counter propagating LH waves generating a whistler wave. Simulation results were presented, based on the LH-whistler mode conversion model of Reference [15], which was generalized by the addition of a nonlinear coupling term. The results of the simulation showed mode-converted whistlers with frequencies near the LH frequency and its harmonic, consistent with the observed spectrum during Exp. 1. It was also shown that the absence of any VLF features near the LH frequency during Exp. 2 was due to the fact that Exp. 2 used short heating pulses. Namely, it was shown that the pulse period was not sufficiently long to develop significant meter-scale striations, thus preventing an efficient linear coupling from LH waves to whistlers. The nonlinear coupling, however, only requires the density fluctuations of another LH wave, and thus does not directly rely on striations. This is consistent with the whistler spectrum observed during Exp. 2. The mode conversion mechanisms have implications on VLF generation with subsequent injection to the radiation belts in order to trigger particle precipitation; the mechanisms could be a source for VLF generation in regions where the electrojet is absent.

Appendix A

The Kinetic and Fluid Descriptions of Plasmas

The following discussion of the kinetic description of plasma physics and the subsequent derivation of the plasma fluid equations is meant to be a handy reference for those who may not be as initiated as a plasma physics practitioner, such as a curious graduate student or one that has just began their studies in this field.

A.1 The Boltzmann Equation

It was mentioned in the beginning of Section 1.2.2 that the intracability of the brute force multi-particle description of a plasma must be superseded by statistical mechanics, an approach to describing many-particle systems with a *distribution function*. The theoretical foundations of statistical mechanics rely on the important notion of a *phase space* - a 6 dimensional space representing the state of a particle, corresponding to the components of its position and momentum (or velocity). The *phase space density* of a single particle with position $\mathbf{x}_1(t)$ and velocity $\mathbf{v}_1(t)$ is simply $\delta^3(\mathbf{x} - \mathbf{x}_1(t))\delta^3(\mathbf{v} - \mathbf{v}_1(t))$. A system of N particles, with positions $\mathbf{X}_k(t)$ and velocities $\mathbf{V}_k(t)$, can be represented by a $6N$ dimensional phase space density,

or distribution function (f_N), given by

$$\frac{df_N}{dt} = \frac{\partial f_N}{\partial t} + \sum_k \mathbf{v}_k \cdot \frac{\partial}{\partial \mathbf{x}_k} f_N + \sum_k \mathbf{a}_k \cdot \frac{\partial}{\partial \mathbf{v}_k} f_N = 0. \tag{A.1}$$

At this stage the problem is still intractable since N is very large. Fortunately, the complexity of the $6N$ dimensional equation (A.1) can be reduced by taking integrals of the distribution function over the individual phase spaces[1] of particles contributing to f_N. A single such integrating gives a *reduced distribution function*, $f_{N-1}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{N-1}, \mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_{N-1}, t)$, where the number of independent coordinates has been reduced by one. Consecutive integrations over the phase space of each particle give a k^{th} reduced distribution functions, f_k, defined by

$$f_k(\mathbf{x}_1, ..., \mathbf{x}_k, \mathbf{v}_1, ..., \mathbf{v}_k, t) = \frac{1}{V^{N-k}} \int d\mathbf{x}_{k+1} d\mathbf{v}_{k+1} ... d\mathbf{x}_N d\mathbf{v}_N \; f_N(\mathbf{x}_1, ..., \mathbf{x}_N, \mathbf{v}_1, ..., \mathbf{v}_N, t), \tag{A.2}$$

where V is a finite spatial volume where f_N is nonzero. Performing $N-1$ integrations of f_N gives the single particle distribution function[2] $f_1(\mathbf{x}_1, \mathbf{v}_1, t)$, and its evolution equation can be obtained by performing $N-1$ integrations of (A.1).

This reduction procedure - the so-called *Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy* [55, Ch. 2] - may sound straightforward, but in practice is not since integrations over subspaces of the phase space produce non-vanishing

[1] Such an integration, for instance, with respect to the whole phase space of particle k, gets rid of the coordinates of this particle and thus its individuality. Its contribution to f_N and its dynamics is smeared out to all the remaining particles as an average effect on the distribution function. The loss of individuality is no problem in plasma physics because all electrons are indistinguishable as are all ions of the same kind.

[2] This reduced distribution function depends merely on the phase space coordinates of one particle, which can be any electron or ion in the plasma, and describes all electrons or ions equally well, distinguishing them only with respect to their velocity and positions.

correlation terms. The correlation terms are often lumped into one term (C_α) on the Right Hand Side (RHS) of the equation. The equation that finally emerges is the *Boltzmann equation* for the *distribution function*[3] $f_\alpha(\mathbf{x}, \mathbf{v}, t)$ given by

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{x}} f_\alpha + \frac{\mathbf{F}_\alpha}{m_\alpha} \cdot \frac{\partial}{\partial \mathbf{v}} f_\alpha = C_\alpha, \tag{A.3}$$

where the subscript α indexes the species that comprise the plasma (e.g. $\alpha = e, i$), $\mathbf{F}_\alpha$ is the corresponding force, m_α is the mass of one particle of the species, and C_α is the so-called *collision operator*: a combined effect of collision of species α with itself all other species in the system (including neutrals, if it is a partially ionized plasma like the ionosphere); this is made more transparent by writing C_α as a sum over the contributions: $C_\alpha = \sum_\beta C_{\alpha\beta}(f_\alpha, f_\beta)$, where $C_{\alpha\beta}(f_\alpha, f_\beta)$ gives the change per unit time in the distribution function for particles of species α due to collisions with particles of species β, and must have the following properties [8] if collisions between particles respectively conserve particle number, momentum and energy:

$$\int C_{\alpha\beta}\, d^3v = 0; \quad \int m_\alpha \mathbf{v} C_{\alpha\alpha}\, d^3v = 0; \quad \int \frac{1}{2} m_\alpha v^2 C_{\alpha\alpha}\, d^3v = 0. \tag{A.4}$$

For a plasma $\mathbf{F}_\alpha$ is given by the Lorentz force, which upon plugging into (A.3) gives the form of the Boltzmann equation that is undoubtedly the cornerstone of

[3]The term *distribution function* in most textbooks and literature refers to the reduced distribution function $f(\mathbf{x}, \mathbf{v}, t)$, and for brevity will henceforth be done here as well. Also, since this procedure is true for any one particle, the subscript labeling for $\mathbf{x}_1, \mathbf{v}_1$ or f_1 has been dropped.

plasma kinetic theory:

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha + \frac{q_\alpha}{m_\alpha}(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial}{\partial \mathbf{v}} f_\alpha = \sum_\beta C_{\alpha\beta}(f_\alpha, f_\beta), \tag{A.5}$$

where q_α is the particle charge, and $\mathbf{E}$ and $\mathbf{B}$ are now the macroscopic fields[4]. When particle collisions are very rare for the time scale of interest, C_α can be set to zero and Equation (A.5) is then referred to as the *Vlasov equation*, used to model so-called *collisionless plasmas*.

Since f_α is a phase space number density, the familiar (spatial) number density[5], n_α, is obtained by integrating f_α over all velocity space; and bulk flow velocity, $\mathbf{u}_\alpha$, as one would expect, is given by the average of $\mathbf{v}$:

$$n_\alpha(\mathbf{x}, t) = \int f_\alpha(\mathbf{x}, \mathbf{v}, t)\ d^3v, \tag{A.6a}$$

$$\mathbf{u}_\alpha(\mathbf{x}, t) = \frac{\int \mathbf{v} f_\alpha(\mathbf{x}, \mathbf{v}, t)\ d^3v}{\int f_\alpha(\mathbf{x}, \mathbf{v}, t)\ d^3v} = \frac{1}{n_\alpha} \int \mathbf{v} f_\alpha(\mathbf{x}, \mathbf{v}, t)\ d^3v \equiv \langle \mathbf{v} \rangle_\alpha\,, \tag{A.6b}$$

where $\langle ... \rangle$ denotes average. The evolution of the electric and magnetic fields in (A.5) are still given by Maxwell's equations (1.3), but the charge density and current

[4]Due to the procedure that leads to Equation (A.3), $\mathbf{F}_\alpha$ is a phase space averaged (macroscopic) quantity, as are the associated fields.

[5]For brevity, *density* is used synonymously with *number density* throughout this dissertation due to its common usage in the field. Mass density, on the other hand, is seldom referenced directly, and as such will be referred to explicitly as *mass density* to avoid confusion; the same goes for charge density.

density are now similarly given as integrals:

$$\rho(\mathbf{x},t) = \sum_{\alpha} q_\alpha \int f_\alpha(\mathbf{x},\mathbf{v},t)\ d^3v, \tag{A.7a}$$

$$\mathbf{J}(\mathbf{x},t) = \sum_{\alpha} q_\alpha \int \mathbf{v} f_\alpha(\mathbf{x},\mathbf{v},t)\ d^3v, \tag{A.7b}$$

where the summation is over every charged species α (e.g. $\alpha = e, i$) in the plasma. The set of equations (1.3), (A.5), and (A.7) form a complete, closed and self consistent kinetic description of a plasma.

A.2 The Fluid Picture

This section shows that plasma dynamics can be described by macroscopic variables alone, corresponding to the more familiar fluid description characterized by thermodynamic quantities such as number density (n), flow velocity ($\mathbf{u}$), temperature (T), and pressure (p). Amazingly, the *fluid equations* that describe the evolution and coupling of these quantities can be derived from the Boltzmann equation, if the proper assumptions are invoked. Namely, for a sufficiently collisional plasma the form of the distribution function is approximately known, making it unnecessary to solve for it from first principles. It is a general result of statistical mechanics that the distribution function tends to a *Maxwellian distribution*[6] in thermal equilibrium, characterized by a constant temperature T. This concept can be extended to a local group of particles when the *collision frequency* is sufficiently high, such that those particles will reach a *local thermal equilibrium* characterized

[6] Also called the Boltzmann or Maxwell-Boltzmann distribution.

by a temperature at that position and time, $T(\mathbf{x}, t)$. Thus the distribution function can be approximated by the following *local* Maxwellian distribution:

$$f(\mathbf{x}, \mathbf{v}, t) \approx f_M(\mathbf{x}, \mathbf{v}, t) \equiv \frac{n(\mathbf{x}, t)}{(2\pi T(\mathbf{x}, t)/m)^{3/2}} e^{-\frac{m}{2T(\mathbf{x},t)}(\mathbf{v}-\mathbf{u}(\mathbf{x},t))^2}, \tag{A.8}$$

where the temperature is measured in energy units for convenience. It can be verified that f_M is consistent with Equations (A.6a), (A.6b), and the usual notion of temperature - defined by $\frac{3}{2}T(\mathbf{x}, t) = \frac{1}{2}m \left\langle (\mathbf{v} - \mathbf{u})^2 \right\rangle$. The exact distribution function will be the Maxwellian part, f_M, plus some unknown perturbation, δf:

$$f(\mathbf{x}, \mathbf{v}, t) = f_M(\mathbf{x}, \mathbf{v}, t) + \delta f. \tag{A.9}$$

To proceed further and derive the fluid equations, note that the three equations in (A.4) are in essence the zeroth, first and second *moments* of the collision operator, respectively. The idea now is to take *moments of the Boltzmann equation* with the same weight factors as in (A.4). This process is relatively straightforward, but algebraically quite cumbersome. The zeroth moment will be worked out in detail, since it is the simplest and sets the stage for the rest. The first moment will be worked in less detail, while the third moment will simply be quoted for the sake of brevity and since standard textbooks carry out the procedure [8, 48, 55, 12].

The zeroth moment of the Boltzmann equation is obtained by integrating (A.5) over velocity space:

$$\int \left\{ \frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha + \frac{\mathbf{F}_\alpha}{m_\alpha} \cdot \frac{\partial f_\alpha}{\partial \mathbf{v}} = \sum_\beta C_{\alpha\beta}(f_\alpha, f_\beta) \right\} d^3v, \tag{A.10}$$

where $\mathbf{F}_\alpha$ represents the Lorentz force. The RHS of (A.10) is zero by virtue of (A.4), while the first term on the Left Hand Side (LHS) simplifies since velocity integration goes past the partial time derivative and acts on the distribution function alone, and thus simplifies by the definition of number density (A.6a):

$$\frac{\partial n_\alpha}{\partial t} + \int \mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha \; d^3v + \int \frac{\mathbf{F}_\alpha}{m_\alpha} \cdot \frac{\partial f_\alpha}{\partial \mathbf{v}} \; d^3v = 0. \tag{A.11}$$

Focusing on the third term of the LHS, note that $\partial/\partial\mathbf{v} \cdot (f_\alpha \mathbf{F}_\alpha) = f_\alpha(\partial/\partial\mathbf{v} \cdot \mathbf{F}_\alpha) + \mathbf{F}_\alpha \cdot (\partial f_\alpha/\partial\mathbf{v})$, thus we can integrate by parts and write

$$\begin{aligned}
\int \mathbf{F}_\alpha \cdot \frac{\partial f_\alpha}{\partial \mathbf{v}} \; d^3v &= -\int f_\alpha \; \frac{\partial}{\partial \mathbf{v}} \cdot \mathbf{F}_\alpha \; d^3v + \int \frac{\partial}{\partial \mathbf{v}} \cdot (f_\alpha \mathbf{F}_\alpha) \; d^3v \\
&= -\int f_\alpha \; \frac{\partial}{\partial \mathbf{v}} \cdot \mathbf{F}_\alpha \; d^3v + \int_S f_\alpha \mathbf{F}_\alpha \cdot \hat{\mathbf{n}} \; dS \\
&= -\int f_\alpha \; \frac{\partial}{\partial \mathbf{v}} \cdot \mathbf{F}_\alpha \; d^3v = 0
\end{aligned}$$

where the middle two lines utilized the divergence theorem, and the surface integral vanished since the distribution function is zero at infinity. The last step is because $\partial/\partial\mathbf{v} \cdot \mathbf{F}_\alpha = 0$, which can be proven by recalling that $\mathbf{F}_\alpha$ is the Lorentz force and thus has an explicit velocity dependence *only* through the velocity cross product:

$$\begin{aligned}
\frac{1}{q_\alpha} \; \frac{\partial}{\partial \mathbf{v}} \cdot \mathbf{F}_\alpha &= \frac{\partial}{\partial \mathbf{v}} \cdot (\mathbf{E} + \mathbf{v} \times \mathbf{B}) = \frac{\partial}{\partial \mathbf{v}} \cdot (\mathbf{v} \times \mathbf{B}) \\
&= \mathbf{B} \cdot \left(\frac{\partial}{\partial \mathbf{v}} \times \mathbf{v} \right) - \mathbf{v} \cdot \left(\frac{\partial}{\partial \mathbf{v}} \times \mathbf{B} \right) = 0
\end{aligned}$$

where an identity for the divergence of a curl is used, and the fact that $\mathbf{v}$ has no velocity curl. Finally, noting that $\boldsymbol{\nabla} \cdot (\mathbf{v} f_\alpha) = f_\alpha \boldsymbol{\nabla} \cdot \mathbf{v} + \mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha = \mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha$ (since

$\boldsymbol{\nabla} \cdot \mathbf{v} = 0$), the middle term on the LHS of (A.11) can be evaluated to give

$$\int \mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha \; d^3v = \int \boldsymbol{\nabla} \cdot (\mathbf{v} f_\alpha) \; d^3v = \boldsymbol{\nabla} \cdot \int (\mathbf{v} f_\alpha) \; d^3v = \boldsymbol{\nabla} \cdot (n_\alpha \mathbf{u}_\alpha).$$

Putting the pieces together, the equation that emerges after the algebra is

$$\boxed{\frac{\partial n_\alpha}{\partial t} + \boldsymbol{\nabla} \cdot (n_\alpha \mathbf{u}_\alpha) = 0,} \tag{A.12}$$

and is called the *continuity equation* - the first of three fluid equations.

The first moment of the Boltzmann equation is obtained by multiplying (A.5) with $m_\alpha \mathbf{v}$ and integrating over velocity space:

$$\int m_\alpha \mathbf{v} \left\{ \frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha + \frac{\mathbf{F}_\alpha}{m_\alpha} \cdot \frac{\partial f_\alpha}{\partial \mathbf{v}} = \sum_\beta C_{\alpha\beta}(f_\alpha, f_\beta) \right\} d^3v. \tag{A.13}$$

It is clear that the first term becomes $\int m_\alpha \mathbf{v} \frac{\partial f_\alpha}{\partial t} = \frac{\partial}{\partial t}(m_\alpha n_\alpha \mathbf{u}_\alpha)$. In the second term, note that the product $\mathbf{v}(\mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha)$, which is a vector times a scalar, can be equivalently thought of as $(\mathbf{vv}) \cdot \boldsymbol{\nabla} f_\alpha$, which is a *dyadic tensor*[7] (a matrix) dotted with a vector. With that in mind we can use the identity $\boldsymbol{\nabla} \cdot (\mathbf{vv} f_\alpha) = (\boldsymbol{\nabla} \cdot (\mathbf{vv})) f_\alpha + (\mathbf{vv}) \cdot \boldsymbol{\nabla} f_\alpha$. The first term on the RHS is clearly zero, so the middle term on the LHS of (A.13) becomes $\int m_\alpha \mathbf{vv} \cdot \boldsymbol{\nabla} f_\alpha = \boldsymbol{\nabla} \cdot \int (m_\alpha \mathbf{vv} f_\alpha) \; d^3v = \boldsymbol{\nabla} \cdot (m_\alpha n_\alpha \langle \mathbf{vv} \rangle)$. This can be put into a more useful form by letting $\mathbf{v} = \mathbf{u}_\alpha + \mathbf{w}$, where the "random velocity", $\mathbf{w}$, by definition satisfies $\langle \mathbf{w} \rangle = 0$ since $\langle \mathbf{v} \rangle = \mathbf{u}_\alpha$[8]. It follows that $m_\alpha n_\alpha \langle \mathbf{vv} \rangle =$

[7] A dyadic is a second rank tensor constructed from the outer product of two vectors. In general if we have vectors $\mathbf{a}$ and $\mathbf{b}$, we can construct a dyadic tensor given by $\overline{\overline{\mathbf{T}}} = \mathbf{ab}$ (double over-lines means tensor), whose components are simply given by $T_{ij} = a_i b_j$.

[8] While it should be obvious from the context, for extra clarity this can be written as $\langle \mathbf{v} \rangle_\alpha = \mathbf{u}_\alpha$ since the species dependent f_α is implicit in the average. But this is somewhat cumbersome, so the subscript for the average is henceforth omitted.

$m_\alpha n_\alpha \langle \mathbf{u}_\alpha \mathbf{u}_\alpha + 2\mathbf{u}_\alpha \mathbf{w} + \mathbf{w}\mathbf{w} \rangle = m_\alpha n_\alpha \langle \mathbf{u}_\alpha \mathbf{u}_\alpha + \mathbf{w}\mathbf{w} \rangle = m_\alpha n_\alpha \mathbf{u}_\alpha \mathbf{u}_\alpha + \overline{\overline{\mathbf{P}}}_\alpha$, where $\overline{\overline{\mathbf{P}}}_\alpha \equiv m_\alpha n_\alpha \langle \mathbf{w}\mathbf{w} \rangle$ is the *pressure tensor*. With this definition we obtain $\boldsymbol{\nabla} \cdot (m_\alpha n_\alpha \langle \mathbf{v}\mathbf{v} \rangle) = \boldsymbol{\nabla} \cdot (m_\alpha n_\alpha \mathbf{u}_\alpha \mathbf{u}_\alpha) + \boldsymbol{\nabla} \cdot \mathbf{P}_\alpha$.

Working out the remaining terms and putting everything together we obtain what is effectively the *Navier-Stokes equation* for a plasma, or simply the *momentum equation*, since it describes the dynamics of momentum density:

$$\frac{\partial}{\partial t}(m_\alpha n_\alpha \mathbf{u}_\alpha) + \boldsymbol{\nabla} \cdot (m_\alpha n_\alpha \mathbf{u}_\alpha \mathbf{u}_\alpha) = q_\alpha n_\alpha (\mathbf{E} + \mathbf{u}_\alpha \times \mathbf{B}) - \boldsymbol{\nabla} \cdot \overline{\overline{\mathbf{P}}}_\alpha - \sum_\beta \mathbf{R}_{\alpha\beta},$$

where $\mathbf{R}_{\alpha\beta}$, defined by $\mathbf{R}_{\alpha\beta} \equiv -\int m_\alpha \mathbf{v} C_{\alpha\beta}(f_\alpha, f_\beta)\, d^3v = -\int m_\alpha \mathbf{w} C_{\alpha\beta}(f_\alpha, f_\beta)\, d^3v$, is the "drag" between species α and β. Note that while it is not explicitly stated here, $\mathbf{R}_{\alpha\alpha} = 0$ due to (A.4). The above momentum equation is said to be in "conservative form", but with the use of the continuity equation (A.12) it can be written in its more standard and familiar form:

$$m_\alpha n_\alpha \left(\frac{\partial}{\partial t} + \mathbf{u}_\alpha \cdot \boldsymbol{\nabla} \right) \mathbf{u}_\alpha = q_\alpha n_\alpha (\mathbf{E} + \mathbf{u}_\alpha \times \mathbf{B}) - \boldsymbol{\nabla} \cdot \overline{\overline{\mathbf{P}}}_\alpha - \sum_\beta \mathbf{R}_{\alpha\beta}, \qquad \text{(A.14)}$$

where the differential operator in the parenthesis on the LHS is the so-called *convective derivative*, which represents the rate of change of a quantity in the frame of a fluid element[9]

Note that the zeroth moment, n_α, couples to the first moment, $\mathbf{u}_\alpha$, by the continuity equation (A.12); but, the momentum equation couples $\mathbf{u}_\alpha$ to the second

[9]Indeed, the total rate of change of any space-time dependent quantity, $U(\mathbf{x}(t), t)$, is dU/dt and by the chain rule can be written as $\frac{d}{dt}U(\mathbf{x}(t), t) = \frac{\partial U(\mathbf{x},t)}{\partial t} + \frac{d\mathbf{x}(t)}{dt} \cdot \frac{\partial U(\mathbf{x},t)}{\partial \mathbf{x}} = \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla}\right) U(\mathbf{x}, t)$. For this reason the convective derivative is often written in the compact form $\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla}$.

moment, thus requiring an equation for the pressure tensor $\overline{\overline{\mathbf{P}}}_\alpha$, and so on; there appears to be *no closure* to the moment equations since the equation for each moment depends on the next higher moment. We can break out of this infinite loop by using Equation (A.9); since f_M is known and the remainder (δf) is a perturbation, certain terms are negligible and the moment hierarchy truncates to a finite set of equations. Plugging (A.9) into the definition of the pressure tensor, $\overline{\overline{\mathbf{P}}}_\alpha \equiv m_\alpha n_\alpha \langle \mathbf{ww} \rangle$, we obtain

$$\overline{\overline{\mathbf{P}}}_\alpha = m_\alpha \frac{n_\alpha}{(2\pi T_\alpha/m_\alpha)^{3/2}} \int \mathbf{ww} e^{-\frac{m_\alpha}{2T_\alpha}w^2} d^3w + \delta\overline{\overline{\mathbf{P}}}_\alpha,$$

where $\delta\overline{\overline{\mathbf{P}}}_\alpha$ physically represents the viscous/shear part of the pressure tensor, and is henceforth neglected since it is a small contribution from δf_α. The pressure tensor simplifies since the symmetry and isotropy of the Gaussian distribution forces all off-diagonal components to be zero and each component along the diagonal to be equal:

$$\overline{\overline{\mathbf{P}}}_\alpha = p_\alpha \bar{\bar{\mathbf{I}}};\ p_\alpha = m_\alpha \frac{n_\alpha}{(2\pi T_\alpha/m_\alpha)^{3/2}} \int w_x^2\ e^{-\frac{m_\alpha}{2T_\alpha}w^2} d^3w, \tag{A.15}$$

where $\bar{\bar{\mathbf{I}}}$ is the identity tensor and p_α is the more familiar *scalar pressure*. The expression for the scalar pressure in A.15 can be readily integrated[10] to give

$$p_\alpha = n_\alpha T_\alpha, \tag{A.16}$$

which is nothing more than the *ideal gas law* for each species α (at each point in space and time). Noting that $-\boldsymbol{\nabla} \cdot \overline{\overline{\mathbf{P}}}_\alpha = -\boldsymbol{\nabla} \cdot (p_\alpha \bar{\bar{\mathbf{I}}}) = -\boldsymbol{\nabla} p_\alpha$ we can write (A.14)

[10]This "Gaussian-type" integral can be found in standard integration tables (e.g. [48, Appx. C]). Also, note that the x-component has been chosen arbitrarily; using the y- or z-component will give the same answer since the exponent contains $w^2 = w_x^2 + w_y^2 + w_z^2$ and the differential (velocity) volume is $d^3w = dw_x dw_y dw_z$

in its final form as

$$\boxed{m_\alpha n_\alpha \left(\frac{\partial \mathbf{u}_\alpha}{\partial t} + \mathbf{u}_\alpha \cdot \boldsymbol{\nabla} \mathbf{u}_\alpha \right) = -\boldsymbol{\nabla} p_\alpha + q_\alpha n_\alpha (\mathbf{E} + \mathbf{u}_\alpha \times \mathbf{B}) - \sum_\beta \mathbf{R}_{\alpha\beta}.} \tag{A.17}$$

The drag term can be related to the *momentum transfer collision frequency* between species α and β, $\nu_{\alpha\beta}$, with the expression $\mathbf{R}_{\alpha\beta} = m_\alpha n_\alpha \nu_{\alpha\beta}(\mathbf{u}_\alpha - \mathbf{u}_\beta)$ [48, Ch. 4].

Although the above is a substantial simplification, an equation that expresses the time evolution of p_α (or T_α) is still required to complete the system. This can be obtained from the *second moment* given by

$$\int \frac{1}{2} m_\alpha v^2 \left\{ \frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \boldsymbol{\nabla} f_\alpha + \frac{\mathbf{F}_\alpha}{m_\alpha} \cdot \frac{\partial f_\alpha}{\partial \mathbf{v}} = \sum_\beta C_{\alpha\beta}(f_\alpha, f_\beta) \right\} d^3 v,$$

which can be simplified with an analogous procedure; the details are left out, but can be found in [8, 48]. The end result is the *energy equation*:

$$\frac{\partial}{\partial t} \left(\frac{1}{2} m_\alpha n_\alpha u_\alpha^2 + \frac{3}{2} p_\alpha \right) + \boldsymbol{\nabla} \cdot \left(\frac{1}{2} m_\alpha n_\alpha u_\alpha^2 \mathbf{u}_\alpha + \frac{5}{2} p_\alpha \mathbf{u}_\alpha + \mathbf{q}_\alpha \right) = q_\alpha n_\alpha \mathbf{u}_\alpha \cdot \mathbf{E} - \mathbf{u}_\alpha \cdot \sum_\beta \mathbf{R}_{\alpha\beta} - \sum_\beta Q_{\alpha\beta},$$

where $\mathbf{q}_\alpha$ is the *heat flux vector* defined by $\mathbf{q}_\alpha = \int \frac{1}{2} m_\alpha w^2 \mathbf{w} f_\alpha \, d^3 v = \frac{1}{2} m_\alpha n_\alpha \left\langle w^2 \mathbf{w} \right\rangle$, and the very last term is a sum over terms representing heat exchange between α and β due to collisions, define by $Q_{\alpha\beta} = -\int \frac{1}{2} m_\alpha w^2 C_{\alpha\beta} \, d^3 v$. These heat exchange terms can be complicated, especially if there are inelastic particle collisions (like in the ionosphere) [48, Ch.4,5,9]. Using the continuity and momentum equations, and ideal gas law (A.16), the above conservative form of the energy equation can be

simplified to an equivalent and more familiar form, given by

$$\boxed{\frac{3}{2}n_\alpha\left(\frac{\partial}{\partial t}+\mathbf{u}_\alpha\cdot\boldsymbol{\nabla}\right)T_\alpha+n_\alpha T_\alpha\boldsymbol{\nabla}\cdot\mathbf{u}_\alpha=-\boldsymbol{\nabla}\cdot\mathbf{q}_\alpha-\mathbf{u}_\alpha\cdot\sum_\beta\mathbf{R}_{\alpha\beta}-\sum_\beta Q_{\alpha\beta}.} \tag{A.18}$$

The term $\boldsymbol{\nabla}\cdot\mathbf{q}_\alpha$, which is responsible for thermal conduction in the system, can be discarded from the energy equation if the time scale of the problem is much shorter than the thermal conduction time scale; but, if this is *not* the case, then the *third moment* of the Boltzmann equation involving $\langle w^2\mathbf{w}\rangle$ is required. This gives an equation for $\mathbf{q}_\alpha$, once the proper truncation of all higher order moments is established[11]. The third moment equation is often simplified to an explicit expression for $\mathbf{q}_\alpha$ and subsequently plugged into the energy equation. For example, the simplest and quite widely used expression for heat flux is given by

$$\mathbf{q}_\alpha=-\kappa_\alpha\boldsymbol{\nabla}T_\alpha, \tag{A.19}$$

where κ_α is the *thermal conductivity* and can take on different forms for fully or partially ionized plasmas [48, Ch.5]. The three fluid equations (A.12), (A.17) and (A.18) (with (A.19) plugged in), together with Maxwell's equations (1.3), constitute a mathematically closed and self-consistent fluid description of a plasma.

[11]Details about moments not evaluated here, such as heat conduction and shear stress, can be found in [48] as a part of the "13-moment equations", so-called since there are 13 governing equations: one for density, three for flow velocity, one for pressure, three for the heat flux vector, and five for the (traceless) shear stress tensor.

Appendix B

Some Details About Plasma Waves

B.1 Cold Plasma Dielectric Tensor

The cold plasma dispersion relation is found by Fourier analyzing the linearized system of equations (3.2), first written down in Section 3.1.1; they are comprised of the momentum equation for species α, the Ampère-Maxwell law, Faraday's law, and the quasineutrality condition:

$$\frac{\partial \mathbf{u}_{\alpha,1}}{\partial t} = \frac{q_\alpha}{m_\alpha}(\mathbf{E}_1 + \mathbf{u}_{\alpha,1} \times \mathbf{B}_0), \tag{B.1a}$$

$$\boldsymbol{\nabla} \times \mathbf{B}_1 = \mu_0 \sum_\alpha q_\alpha n_{\alpha,0} \mathbf{u}_{\alpha,1} + \frac{1}{c^2}\frac{\partial \mathbf{E}_1}{\partial t}, \tag{B.1b}$$

$$\boldsymbol{\nabla} \times \mathbf{E}_1 = -\frac{\partial \mathbf{B}_1}{\partial t}, \tag{B.1c}$$

$$\sum_\alpha q_\alpha n_{\alpha,0} = 0. \tag{B.1d}$$

To this end, we assume a harmonic space and time dependence for each quantity. Letting $\mathbf{f}_1$ be a place holder for any of the dynamical quantities (i.e. $\mathbf{u}_{\alpha,1}$, $\mathbf{E}_1$, $\mathbf{B}_1$), we take $\mathbf{f}_1$ to be of the form $\mathbf{f}_1 = \Re\left(\hat{\mathbf{f}}_1 \exp\left(i(\mathbf{k}\cdot\mathbf{x} - \omega t)\right)\right)$, where $\hat{\mathbf{f}}_1$ is the corresponding complex amplitude and $\Re(.)$ denotes the real part. Taking $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$

and $\mathbf{u}_1 = \Re\left(\hat{\mathbf{u}}_1 \exp\left(i\mathbf{k}\cdot\mathbf{x} - i\omega t\right)\right)$, Equation (B.1a) becomes[1]

$$-i\omega\hat{\mathbf{u}}_{\alpha,1} = \frac{q_\alpha}{m_\alpha}\hat{\mathbf{E}}_1 + \Omega_\alpha \hat{\mathbf{u}}_{\alpha,1}\times\hat{\mathbf{z}}, \tag{B.2}$$

where $\Omega_\alpha \equiv q_\alpha B_0/m_\alpha$ is the *algebraic* gyrofrequency for species α; (that is, Ω_α carries the sign of the charge.) Letting $\hat{\mathbf{u}}_{\alpha,1} = \hat{u}^{\parallel}_{\alpha,1}\hat{\mathbf{z}} + \hat{\mathbf{u}}^{\perp}_{\alpha,1}$ we can break up Equation (B.2) into parallel and perpendicular components:

$$\hat{u}^{\parallel}_{\alpha,1} = \frac{i}{\omega}\frac{q_\alpha}{m_\alpha}\hat{E}^{\parallel}_1, \tag{B.3a}$$

$$\hat{\mathbf{u}}^{\perp}_{\alpha,1} = \frac{i}{\omega}\left(\frac{q_\alpha}{m_\alpha}\hat{\mathbf{E}}^{\perp}_1 + \Omega_\alpha\hat{\mathbf{u}}^{\perp}_{\alpha,1}\times\hat{\mathbf{z}}\right). \tag{B.3b}$$

The Ampère-Maxwell equation (B.1b) now takes the complex form

$$i\mathbf{k}\times\hat{\mathbf{B}}_1 = \mu_0\sum_\alpha q_\alpha n_{\alpha,0}\hat{\mathbf{u}}_{\alpha,1} - i\omega\frac{1}{c^2}\hat{\mathbf{E}}_1 = -i\omega\frac{1}{c^2}\left(\frac{1}{-i\omega\epsilon_0}\sum_\alpha q_\alpha n_{\alpha,0}\hat{\mathbf{u}}_{\alpha,1} + \hat{\mathbf{E}}_1\right). \tag{B.4}$$

Using the components of $\mathbf{u}_{\alpha,1}$ from Equation (B.3) and defining the plasma frequency (squared), $\omega^2_{p\alpha} \equiv q^2_\alpha n_{\alpha,0}/(\epsilon_0 m_\alpha)$, the complex current density $\hat{\mathbf{J}}_1 = \sum_\alpha q_\alpha n_\alpha \hat{\mathbf{u}}_{\alpha,1}$ in Equation (B.4) becomes

$$\hat{\mathbf{J}}_1 = -i\omega\epsilon_0\sum_\alpha \omega^2_{p\alpha}\left[\frac{1}{\omega^2-\Omega^2_\alpha}\left(-\hat{\mathbf{E}}^{\perp}_1 + \frac{\Omega_\alpha}{i\omega}\hat{\mathbf{E}}^{\perp}_1\times\hat{\mathbf{z}}\right) - \frac{1}{\omega^2}\hat{E}^{\parallel}_1\right]. \tag{B.5}$$

[1]Note that the "hat" notation is now being used for both the complex amplitudes and the unit vectors, such as $\hat{\mathbf{z}}$. However, no confusion should arise from this dual notation if the context is taken into consideration.

Plugging this expression for $\hat{\mathbf{J}}_1$ into Equation (B.4) gives

$$i\mathbf{k}\times\hat{\mathbf{B}}_1 = -i\omega\mu_0\epsilon_0\left[\hat{\mathbf{E}}_1 - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega^2-\Omega_\alpha^2}\hat{\mathbf{E}}_1^\perp - i\sum_\alpha \frac{\Omega_\alpha}{\omega}\frac{\omega_{p\alpha}^2}{\omega^2-\Omega_\alpha^2}\hat{\mathbf{E}}_1^\perp\times\hat{\mathbf{z}} - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega^2}\hat{E}_1^\parallel\hat{\mathbf{z}}\right].$$

Since $\hat{\mathbf{E}}_1 = \hat{\mathbf{E}}_1^\perp + \hat{E}_1^\parallel\hat{\mathbf{z}}$, the above can be compactly expressed as

$$i\mathbf{k}\times\hat{\mathbf{B}}_1 = -i\omega\mu_0\epsilon_0\left(S\hat{\mathbf{E}}_1^\perp - iD\hat{\mathbf{E}}_1^\perp\times\hat{\mathbf{z}} + P\hat{E}_1^\parallel\hat{\mathbf{z}}\right), \tag{B.6}$$

where S, D, P and convenient quantities $R = S + D$ and $L = S - D$ are defined as

$$S \equiv \frac{1}{2}(R+L) = 1 - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega^2-\Omega_\alpha^2}, \tag{B.7a}$$

$$D \equiv \frac{1}{2}(R-L) = \sum_\alpha \frac{\Omega_\alpha}{\omega}\frac{\omega_{p\alpha}^2}{\omega^2-\Omega_\alpha^2}, \tag{B.7b}$$

$$P \equiv 1 - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega^2}, \tag{B.7c}$$

$$R \equiv 1 - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega(\omega+\Omega_\alpha)}; \quad L \equiv 1 - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega(\omega-\Omega_\alpha)}, \tag{B.7d}$$

The RHS of Equation (B.6) is represented more cleanly as $\overline{\overline{\boldsymbol{\mathcal{E}}}}\cdot\hat{\mathbf{E}} \equiv S\hat{\mathbf{E}}_1^\perp - iD\hat{\mathbf{E}}_1^\perp\times\hat{\mathbf{z}} + P\hat{E}_1^\parallel\hat{\mathbf{z}}$, where the currently unknown tensor $\overline{\overline{\boldsymbol{\mathcal{E}}}}$ can be deduced by breaking up the expression above into the usual Cartesian components[2] $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$:

$$\overline{\overline{\boldsymbol{\mathcal{E}}}}\cdot\hat{\mathbf{E}} = \begin{pmatrix} S\hat{E}_x - iD\hat{E}_y \\ iD\hat{E}_x + S\hat{E}_y \\ P\hat{E}_z \end{pmatrix} = \begin{pmatrix} S & -iD & 0 \\ iD & S & 0 \\ 0 & 0 & P \end{pmatrix}\begin{pmatrix} \hat{E}_x \\ \hat{E}_y \\ \hat{E}_z \end{pmatrix}, \tag{B.8}$$

[2]For convenience the subscript 1 is omitted and the electric field components are written as $\hat{\mathbf{E}}_\perp = \hat{E}_x\hat{\mathbf{x}} + \hat{E}_y\hat{\mathbf{y}}$, and $\hat{E}_\parallel = \hat{E}_z$.

from which we identify the matrix on the RHS as being $\overline{\overline{\boldsymbol{\mathcal{E}}}}$, i.e. the (dimensionless) *cold plasma dielectric tensor*. Equation (B.6) is now written as

$$i\mathbf{k} \times \hat{\mathbf{B}} = -i\omega\mu_0\epsilon_0 \overline{\overline{\boldsymbol{\mathcal{E}}}} \cdot \hat{\mathbf{E}}. \tag{B.9}$$

To make the role of $\overline{\overline{\boldsymbol{\mathcal{E}}}}$ as a dielectric tensor more explicit, note that if we "put back" the derivatives and space-time dependence of the fields in Equation (B.9), and define the effective (dimensional) dielectric tensor $\overline{\overline{\boldsymbol{\epsilon}}} \equiv \epsilon_0 \overline{\overline{\boldsymbol{\mathcal{E}}}}$, we obtain

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \overline{\overline{\boldsymbol{\epsilon}}} \cdot \frac{\partial \mathbf{E}}{\partial t}. \tag{B.10}$$

Or by defining the usual electric displacement ($\mathbf{D} = \overline{\overline{\boldsymbol{\epsilon}}} \cdot \mathbf{E}$) and H-field ($\mathbf{H} = \mathbf{B}/\mu_0$) we obtain the "matter" form of the Ampère-Maxwell equation: $\boldsymbol{\nabla} \times \mathbf{H} = \partial\,\mathbf{D}/\,\partial\,t$; this is why we refer to $\overline{\overline{\boldsymbol{\mathcal{E}}}}$ (or $\overline{\overline{\boldsymbol{\epsilon}}}$) as a dielectric tensor. Taking a time derivative of (B.10) and using Faraday's law (B.1c) gives a wave equation[3] for the electric field:

$$-\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E} = \mu_0 \overline{\overline{\boldsymbol{\epsilon}}} \cdot \frac{\partial^2 \mathbf{E}}{\partial\, t^2}. \tag{B.11}$$

(In Equation (B.11) and elsewhere in this thesis, $\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}$ is written with the understanding that it is shorthand for $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{E})$.)

Going back to Fourier space, this wave equation becomes $\mathbf{k} \times (\mathbf{k} \times \mathbf{E}) =$

[3]Recall that we "put back" the derivatives to get to this wave equation. While putting back the spatial derivatives is fine, strictly speaking the time dependence must be harmonic, which means what we really have is a *Helmholtz equation*. At the end of the day, rigorously working out the type of equation will require uncoupling the original system of equations (if possible) *without* going into Fourier space. Nevertheless, the point being made by writing down Equation (B.11) still stands: the system supports wave-like solutions, but with an effective dielectric tensor rather than the usual ϵ_0.

$-\omega^2\mu_0\epsilon_0\overline{\overline{\boldsymbol{\mathcal{E}}}}\cdot\mathbf{E}$, or more compactly $\boldsymbol{\eta}\times(\boldsymbol{\eta}\times\mathbf{E})+\overline{\overline{\boldsymbol{\mathcal{E}}}}\cdot\mathbf{E}=0$, where

$$\boldsymbol{\eta}=\frac{\mathbf{k}c}{\omega}$$

is the *index of refraction* vector; it points in the direction of propagation, as $\mathbf{k}$ does, but its magnitude is the index of refraction. Using a double cross product identity, $\boldsymbol{\eta}\times(\boldsymbol{\eta}\times\mathbf{E})=-\eta^2\mathbf{E}+\boldsymbol{\eta}(\boldsymbol{\eta}\cdot\mathbf{E})$, the above can be put into a more workable form:

$$\left[\eta^2\overline{\overline{\mathbf{I}}}-\boldsymbol{\eta}\boldsymbol{\eta}-\overline{\overline{\boldsymbol{\mathcal{E}}}}\right]\cdot\mathbf{E}=0, \tag{B.12}$$

where $\overline{\overline{\mathbf{I}}}$ is the identity tensor and $\boldsymbol{\eta}\boldsymbol{\eta}$ is a dyadic tensor defined the usual way. Due to our choice of $\mathbf{B}_0$ pointing in the z-direction and the fact that the plasma has cylindrical symmetry about $\mathbf{B}_0$, we can (without loss of generality) *choose* $\boldsymbol{\eta}$ to lie in the x-z plane. For convenience $\boldsymbol{\eta}$ can be express as $\boldsymbol{\eta}=n_x\hat{\mathbf{x}}+n_z\hat{\mathbf{z}}=\eta\sin\theta\hat{\mathbf{x}}+\eta\cos\theta\hat{\mathbf{z}}$, where $\eta=|\boldsymbol{\eta}|$ and θ is the usual angle between $\mathbf{k}$ and $\mathbf{B}_0$. Using this form of $\boldsymbol{\eta}$, the matrix form of Equation (B.12) can be readily worked out to be

$$\begin{pmatrix} S-\eta^2\cos^2\theta & -iD & \eta^2\cos\theta\sin\theta \\ iD & S-\eta^2 & 0 \\ \eta^2\cos\theta\sin\theta & 0 & P-\eta^2\sin^2\theta \end{pmatrix}\begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}=0. \tag{B.13}$$

In order to have a nontrivial solution for the fields, the determinant of the tensor in Equation (B.13) must vanish. This condition gives the *cold plasma dispersion relation* (CPDR) [54],

$$A\eta^4-B\eta^2+C=0, \tag{B.14}$$

where $A = S\sin^2\theta + P\cos^2\theta$, $B = RL\sin^2\theta + PS(1+\cos^2\theta)$, and $C = PRL$. The solution to Equation (B.14) can be written as a quadratic in η:

$$\eta^2 = \frac{B \pm F}{2A}, \tag{B.15}$$

where $F^2 \equiv B^2 - 4AC$ and may be written in the form $F^2 = (RL - PS)^2\sin^4\theta + 4P^2D^2\cos^2\theta$ [54]. Moreover, the CPDR in the form (B.15) can be recast into a different but equivalent form given by [54]:

$$\tan^2\theta = \frac{-P(\eta^2 - R)(\eta^2 - L)}{(S\eta^2 - RL)(\eta^2 - P)}. \tag{B.16}$$

This form of the CPDR is convenient since the dispersion relations for the various wave mode can be read off easily for the special cases of parallel ($\theta = 0$) and perpendicular ($\theta = \pi/2$) propagation.

Incidentally, the Apptleton-Hartree equation (1.14) from Chapter 1 can be derived from this formalism; the only difference is collisions are neglected here. The work of Appleton and others included electron motion but assumed that ion motion was negligible, which is valid for HF wave propagation in the ionosphere since the HF period is too short for the (much heavier) ions to be able to respond. Based on the formalism above, this simply amounts to taking the sums in S, D, P, R, and L to be only over the electrons. For example, for S we would have

$$S = 1 - \frac{\omega_{pe}^2}{\omega^2 - \Omega_e^2} = 1 - \frac{X}{1 - Y^2},$$

where the X, Y notation has been introduced; the rest can be recast in terms of X and Y in a similar way. Rewriting Equation (B.15) in the form

$$\eta^2 = 1 - \frac{2(A - B + C)}{2A - B \pm \sqrt{B^2 - 4AC}},$$

plugging in A, B, and C in terms of X and Y and simplifying, yields the previously quoted Appleton-Hartree equation (1.14) (with $Z = 0$).

B.2 Evolution Equations for Alfvén and Magnetosonic Waves

It would be nice to have a toy picture of Alfvén and MS wave propagation, and to this end we provide here the derivation of the evolution equations for the electric field and vector potential in the MHD limit. For simplicity we assume one (singly ionized) ion species, a constant vertical background magnetic field $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$, constant background plasma density n_0, no collisions, and a low-β plasma so that the pressure terms can be neglected.

With these assumptions, the electron and ion linearized momentum equations are

$$m_e n_0 \frac{\partial \mathbf{v}_e}{\partial t} = -en_0(\mathbf{E} + \mathbf{v}_e \times \mathbf{B}_0), \quad m_i n_0 \frac{\partial \mathbf{v}_i}{\partial t} = en_0(\mathbf{E} + \mathbf{v}_i \times \mathbf{B}_0),$$

where we have now switched the labeling of flow velocity from $\mathbf{u}$ to $\mathbf{v}$. By adding the above electron and momentum equations together, and defining the average velocity $\mathbf{v} \equiv (m_e\mathbf{v}_e + m_i\mathbf{v}_i)/(m_e + m_i)$ and current density $\mathbf{J} = en_0(\mathbf{v}_i - \mathbf{v}_e)$, we obtain

$$mn_0 \frac{\partial \mathbf{v}}{\partial t} = \mathbf{J} \times \mathbf{B}_0, \tag{B.17}$$

where $m = m_e + m_i \approx m_i$ is the total mass. This is effectively the momentum equation for a single conducting fluid (in the low frequency range). Note that the single fluid does not "see" the electric field since the electrons and ions move together. Similarly, by multiplying the electron momentum equation by m_i and the ion momentum equation by m_e and subtracting them we obtain

$$\mathbf{E} + \mathbf{v} \times \mathbf{B}_0 = \frac{1}{emn_0} \left(\frac{m_e m_i}{e} \frac{\partial \mathbf{J}}{\partial t} + (m_i - m_e) \mathbf{J} \times \mathbf{B}_0 \right).$$

This is the so-called *generalized Ohm's law*, since it is an explicit relationship between the electric field and quantities like the fluid velocity and current density. The first term on the RHS is the "inertial" term and can be neglected for low frequencies, while the second term on the RHS, the so-called *Hall current* term, can be neglected for frequencies much below the ion gyrofrequency. With these simplifications we are left with the ideal Ohm's law

$$\mathbf{E} + \mathbf{v} \times \mathbf{B}_0 = 0. \tag{B.18}$$

The final two necessary equations are Faraday's law, $\partial \mathbf{B} / \partial t = -\boldsymbol{\nabla} \times \mathbf{E}$, and Ampère's law $\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \mathbf{J}$, where the displacement current has been neglected since we are dealing with low frequencies. Now, we can take a time derivative of (B.18), use Equation (B.17) to eliminate $\partial \mathbf{v} / \partial t$, and then use Ampère's law to eliminate $\mathbf{J}$; working out the algebra gives

$$\frac{\partial \mathbf{E}}{\partial t} = -\frac{1}{\mu_0 m n_0} \left[(\boldsymbol{\nabla} \times \mathbf{B}) \times \mathbf{B}_0 \right] \times \mathbf{B}_0 = -v_A^2 \left[(\boldsymbol{\nabla} \times \mathbf{B}) \times \hat{\mathbf{z}} \right] \times \hat{\mathbf{z}}, \tag{B.19}$$

where $v_A = B_0/\sqrt{\mu_0 m n_0}$ is the Alfvén speed. Along with Faraday's law, this equation can be turned into a wave equation:

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = v_A^2 \left[(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}) \times \hat{\mathbf{z}} \right] \times \hat{\mathbf{z}}. \tag{B.20}$$

The double cross product is effectively the inverse dielectric tensor. The SA and MS waves are contained within this equation, which can be shown by Fourier analyzing Equation (B.20) and finding the dispersion relations. For convenience we introduce the vector potential $\mathbf{A}$ such that $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$; Faraday's law tells us that $\mathbf{E} = -\partial \mathbf{A}/\partial t - \boldsymbol{\nabla}\phi$, where ϕ is the scalar potential. Choosing the gauge $\phi = 0$ we are left with the following system of equations:

$$\frac{\partial \mathbf{E}}{\partial t} = -v_A^2 \left[(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{A}) \times \hat{\mathbf{z}} \right] \times \hat{\mathbf{z}}, \tag{B.21a}$$

$$\frac{\partial \mathbf{A}}{\partial t} = -\mathbf{E}. \tag{B.21b}$$

Appendix C

Governing Equations of ICD

The purpose here is to highlight the steps that derive the ICD equations (3.22), starting with the basic equations (3.21). The general procedure is to define convenient quantities in an effort to make things more compact, and subsequently decoupled the basic equations via various combinations of differentiation and substitution.

For convenience Reference [14] defines $\Gamma_{en} \equiv \nu_{en}/\omega_{ce}$ and $\Gamma_{in} \equiv \nu_{in}/\omega_{ci}$, where $\omega_{ce} = eB_0/m_e$ and $\omega_{ci} = eB_0/m_i$ are the (positive) gyrofrequencies; subsequently defining the tensors $\overline{\overline{\mathbf{R}}}_e$ and $\overline{\overline{\mathbf{R}}}_i$

$$\overline{\overline{\mathbf{R}}}_e \mathbf{v}_e \equiv \frac{\mathbf{v}_e \times \mathbf{B}_0}{B_0} + \Gamma_{en}\mathbf{v}_e, \tag{C.1}$$

$$\overline{\overline{\mathbf{R}}}_i \mathbf{v}_i \equiv \frac{\mathbf{v}_i \times \mathbf{B}_0}{B_0} - \Gamma_{in}\mathbf{v}_i, \tag{C.2}$$

the electron and ion momentum equations (3.21c), (3.21d) can be written more compactly as

$$\mathbf{E} = -B_0 \overline{\overline{\mathbf{R}}}_e \mathbf{v}_e - \frac{\boldsymbol{\nabla} p_e}{e n_0}, \tag{C.3}$$

$$\frac{\partial \mathbf{v}_i}{\partial t} = \frac{e}{m_i}(\mathbf{E} + B_0 \overline{\overline{\mathbf{R}}}_i \mathbf{v}_i). \tag{C.4}$$

Taking a time derivative of Equation (C.3),

$$\frac{\partial \mathbf{E}}{\partial t} = -B_0 \overline{\overline{\mathbf{R}}}_e \frac{\partial \mathbf{v}_e}{\partial t} - \frac{1}{en_0} \boldsymbol{\nabla} \frac{\partial p_e}{\partial t},$$

and using Ampère's law (3.21b), $\mathbf{v}_e = \mathbf{v}_i - \boldsymbol{\nabla} \times \mathbf{B}/(\mu_0 e n_0)$, to plug in for v_e gives

$$\frac{\partial \mathbf{E}}{\partial t} = -B_0 \overline{\overline{\mathbf{R}}}_e \left(\frac{\partial \mathbf{v}_i}{\partial t} - \frac{1}{\mu_0 e n_0} \boldsymbol{\nabla} \times \frac{\partial \mathbf{B}}{\partial t} \right) - \frac{1}{en_0} \boldsymbol{\nabla} \frac{\partial p_e}{\partial t}.$$

Plugging in for $\partial \mathbf{v}_i / \partial t$ by using Equation (C.4), the previous equation becomes

$$\frac{\partial \mathbf{E}}{\partial t} = -B_0 \overline{\overline{\mathbf{R}}}_e \left(\frac{e}{m_i} (\mathbf{E} + B_0 \overline{\overline{\mathbf{R}}}_i \mathbf{v}_i) - \frac{1}{\mu_0 e n_0} \boldsymbol{\nabla} \times \frac{\partial \mathbf{B}}{\partial t} \right) - \frac{1}{en_0} \boldsymbol{\nabla} \frac{\partial p_e}{\partial t}. \tag{C.5}$$

The $\mathbf{v}_i$ in Equation (C.5) can be related to $\mathbf{B}$ by using Ampère's law, $\mathbf{v}_i = \mathbf{v}_e + \boldsymbol{\nabla} \times \mathbf{B}/(\mu_0 e n_0)$; meanwhile $\mathbf{v}_e$ can be expressed in terms of $\mathbf{E}$ by using Equation (C.3): $\mathbf{v}_e = -B_0^{-1} \overline{\overline{\mathbf{R}}}_e^{-1} (\mathbf{E} + \boldsymbol{\nabla} p_e / e n_0)$. Putting these together, $\mathbf{v}_i$ can now be expressed as

$$\mathbf{v}_i = \frac{\boldsymbol{\nabla} \times \mathbf{B}}{\mu_0 e n_0} - \frac{1}{B_0} \overline{\overline{\mathbf{R}}}_e^{-1} \left(\mathbf{E} + \frac{\boldsymbol{\nabla} p_e}{e n_0} \right),$$

which upon plugging into Equation (C.5) yields an equation in terms of the electric and magnetic fields alone:

$$\frac{\partial \mathbf{E}}{\partial t} = -B_0 \overline{\overline{\mathbf{R}}}_e \left[\frac{e}{m_i} \left(\mathbf{E} + B_0 \overline{\overline{\mathbf{R}}}_i \left(\frac{\boldsymbol{\nabla} \times \mathbf{B}}{\mu_0 e n_0} - \frac{1}{B_0} \overline{\overline{\mathbf{R}}}_e^{-1} \left(\mathbf{E} + \frac{\boldsymbol{\nabla} p_e}{e n_0} \right) \right) \right) - \frac{1}{\mu_0 e n_0} \boldsymbol{\nabla} \times \frac{\partial \mathbf{B}}{\partial t} \right] - \frac{1}{en_0} \boldsymbol{\nabla} \frac{\partial p_e}{\partial t}.$$

Noting that the $\overline{\overline{\mathbf{R}}}_e$ and $\overline{\overline{\mathbf{R}}}_i$ tensors *commute*[1] (i.e. $\overline{\overline{\mathbf{R}}}_e\overline{\overline{\mathbf{R}}}_i = \overline{\overline{\mathbf{R}}}_i\overline{\overline{\mathbf{R}}}_e$) and using Faraday's law (3.21a), the above simplifies to

$$\frac{\partial \mathbf{E}}{\partial t} = -\omega_{ci}(\overline{\overline{\mathbf{R}}}_e - \overline{\overline{\mathbf{R}}}_i)\mathbf{E} - \frac{B_0^2}{\mu_0 m_i n_0}\overline{\overline{\mathbf{R}}}_e\overline{\overline{\mathbf{R}}}_i\boldsymbol{\nabla}\times\mathbf{B} - \frac{B_0}{\mu_0 e n_0}\overline{\overline{\mathbf{R}}}_e\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{E} - \frac{1}{e n_0}\boldsymbol{\nabla}\frac{\partial p_e}{\partial t} + \omega_{ci}\overline{\overline{\mathbf{R}}}_i\frac{\boldsymbol{\nabla} p_e}{e n_0}. \tag{C.6}$$

It is clear that the factor $B_0^2/(\mu_0 m_i n_0) \equiv v_A^2$ is the Alfvén speed squared. Moreover, as in Equation (B.10) we define the dielectric tensor, $\overline{\overline{\boldsymbol{\epsilon}}}$, and its inverse as

$$\overline{\overline{\boldsymbol{\epsilon}}} = -\frac{\epsilon_0 c^2}{v_A^2}(\overline{\overline{\mathbf{R}}}_e\overline{\overline{\mathbf{R}}}_i)^{-1}, \quad \overline{\overline{\boldsymbol{\epsilon}}}^{-1} = -\frac{v_A^2}{\epsilon_0 c^2}\overline{\overline{\mathbf{R}}}_e\overline{\overline{\mathbf{R}}}_i.$$

With these definitions the above can be written as

$$\frac{\partial \mathbf{E}}{\partial t} = \overline{\overline{\boldsymbol{\epsilon}}}^{-1}\left(-\overline{\overline{\boldsymbol{\sigma}}}\mathbf{E} + \frac{\boldsymbol{\nabla}\times\mathbf{B}}{\mu_0}\right) - \frac{B_0}{\mu_0 e n_0}\overline{\overline{\mathbf{R}}}_e\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{E} + \left(\omega_{ci}\overline{\overline{\mathbf{R}}}_i - \frac{\partial}{\partial\, t}\right)\frac{\boldsymbol{\nabla} p_e}{e n_0}, \tag{C.7}$$

where the conductivity tensor $\overline{\overline{\boldsymbol{\sigma}}}$ has been identified and is defined by $\overline{\overline{\boldsymbol{\epsilon}}}^{-1}\overline{\overline{\boldsymbol{\sigma}}} = \omega_{ci}(\overline{\overline{\mathbf{R}}}_e - \overline{\overline{\mathbf{R}}}_i)$. Multiplying through by $\overline{\overline{\boldsymbol{\epsilon}}}$ gives the expression for the conductivity tensor:

$$\overline{\overline{\boldsymbol{\sigma}}} = \omega_{ci}\frac{\epsilon_0 c^2}{v_A^2}(\overline{\overline{\mathbf{R}}}_e\overline{\overline{\mathbf{R}}}_i)^{-1}(\overline{\overline{\mathbf{R}}}_i - \overline{\overline{\mathbf{R}}}_e) = \frac{\omega_{ci}\epsilon_0 c^2}{v_A^2}(\overline{\overline{\mathbf{R}}}_e^{-1} - \overline{\overline{\mathbf{R}}}_i^{-1}) = \tilde{\sigma}(\overline{\overline{\mathbf{R}}}_e^{-1} - \overline{\overline{\mathbf{R}}}_i^{-1}),$$

where the definition $\omega_{ci}\epsilon_0 c^2/v_A^2 = e n_0/B_0 \equiv \tilde{\sigma}$ has been made. Note that expressions containing the difference $\overline{\overline{\mathbf{R}}}_e - \overline{\overline{\mathbf{R}}}_i$ (e.g. in Equation (C.6)) simplify to $(\overline{\overline{\mathbf{R}}}_e - \overline{\overline{\mathbf{R}}}_i)\mathbf{E} = (\Gamma_{en} + \Gamma_{in})\mathbf{E}$, since the cross product contributions cancel. With this simplification

[1] This follows directly from the definitions of $\overline{\overline{\mathbf{R}}}_e$ and $\overline{\overline{\mathbf{R}}}_i$: letting $\hat{\mathbf{b}}_0 \equiv \mathbf{B}_0/B_0$ and $\mathbf{F}$ be an arbitrary function, we have $\overline{\overline{\mathbf{R}}}_e\overline{\overline{\mathbf{R}}}_i\mathbf{F} = (\mathbf{F}\times\hat{\mathbf{b}}_0 - \Gamma_{in}\mathbf{F})\times\hat{\mathbf{b}}_0 + \Gamma_{en}(\mathbf{F}\times\hat{\mathbf{b}}_0 - \Gamma_{in}\mathbf{F}) = (\mathbf{F}\times\hat{\mathbf{b}}_0)\times\hat{\mathbf{b}}_0 - \Gamma_{in}\mathbf{F}\times\hat{\mathbf{b}}_0 + \Gamma_{en}\mathbf{F}\times\hat{\mathbf{b}}_0 - \Gamma_{en}\Gamma_{in}\mathbf{F} = (\mathbf{F}\times\hat{\mathbf{b}}_0 + \Gamma_{en}\mathbf{F})\times\hat{\mathbf{b}}_0 - \Gamma_{in}(\mathbf{F}\times\hat{\mathbf{b}}_0 + \Gamma_{en}\mathbf{F}) = \overline{\overline{\mathbf{R}}}_i\overline{\overline{\mathbf{R}}}_e\mathbf{F}$.

and the above definitions, Equation (C.7) reduces to

$$\frac{\partial \mathbf{E}}{\partial t} = -\omega_{ci}(\Gamma_{en} + \Gamma_{in})\mathbf{E} + \frac{\overline{\overline{\boldsymbol{\epsilon}}}^{-1}}{\mu_0}\left[\boldsymbol{\nabla} \times \mathbf{B}\right] - \frac{\overline{\overline{\mathbf{R}}}_e\left[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}\right]}{\mu_0 \tilde{\sigma}} + \left(\omega_{ci}\overline{\overline{\mathbf{R}}}_i - \frac{\partial}{\partial t}\right)\frac{\boldsymbol{\nabla} p_e}{en_0}. \tag{C.8}$$

Faraday's law (3.21a) and Equation (C.8) constitute a closed mathematical system of equation for the $\mathbf{E}$ and $\mathbf{B}$ fields. Nevertheless, it is convenient to introduce the vector potential $\mathbf{A}$, defined by $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$; Faraday's law then reduces to $\mathbf{E} = -\boldsymbol{\nabla}\phi - \partial\,\mathbf{A}/\,\partial\, t$, where ϕ is the scalar potential. By choosing the gauge $\phi = 0$, this reduces further to $\mathbf{E} = -\,\partial\,\mathbf{A}/\,\partial\, t$ and finally yields the following coupled system of partial differential equations for $\mathbf{E}$ and $\mathbf{A}$, first quoted as Equations (3.22):

$$\frac{\partial \mathbf{A}}{\partial t} = -\mathbf{E}, \tag{C.9a}$$

$$\frac{\partial \mathbf{E}}{\partial t} = -\omega_{ci}(\Gamma_{en} + \Gamma_{in})\mathbf{E} + \frac{\overline{\overline{\boldsymbol{\epsilon}}}^{-1}\left[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{A}\right]}{\mu_0} - \frac{\overline{\overline{\mathbf{R}}}_e\left[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}\right]}{\mu_0 \tilde{\sigma}} + \left(\omega_{ci}\overline{\overline{\mathbf{R}}}_i - \frac{\partial}{\partial t}\right)\frac{\boldsymbol{\nabla} p_e}{en_0}. \tag{C.9b}$$

By choosing $\mathbf{B}_0$ to be in, say, the z direction, the matrix form of the tensors can be written down. For example, the $\overline{\overline{\mathbf{R}}}_e$, $\overline{\overline{\mathbf{R}}}_i$, and $\overline{\overline{\boldsymbol{\sigma}}}$ tensors are given by

$$\overline{\overline{\mathbf{R}}}_e = \begin{pmatrix} \Gamma_{en} & 1 & 0 \\ -1 & \Gamma_{en} & 0 \\ 0 & 0 & \Gamma_{en} \end{pmatrix}, \quad \overline{\overline{\mathbf{R}}}_i = \begin{pmatrix} -\Gamma_{in} & 1 & 0 \\ -1 & -\Gamma_{in} & 0 \\ 0 & 0 & -\Gamma_{in} \end{pmatrix}, \quad \overline{\overline{\boldsymbol{\sigma}}} = \begin{pmatrix} \sigma_P & -\sigma_H & 0 \\ \sigma_H & \sigma_P & 0 \\ 0 & 0 & \sigma_\parallel \end{pmatrix}, \tag{C.10}$$

where the Pedersen (σ_P), Hall (σ_H), and parallel ($\sigma_\parallel$) conductivities are

$$\sigma_P = \tilde{\sigma}\left[\frac{\Gamma_{en}}{1+\Gamma_{en}^2} + \frac{\Gamma_{in}}{1+\Gamma_{in}^2}\right], \quad \sigma_H = \tilde{\sigma}\left[\frac{1}{1+\Gamma_{en}^2} - \frac{1}{1+\Gamma_{in}^2}\right], \quad \sigma_\parallel = \tilde{\sigma}\left[\frac{1}{\Gamma_{en}} + \frac{1}{\Gamma_{in}}\right]. \tag{C.11}$$

The inverse of the dielectric tensor is given by

$$\overline{\overline{\boldsymbol{\epsilon}}}^{-1} = \epsilon_0^{-1}\frac{v_A^2}{c^2}\begin{pmatrix} 1+\Gamma_{in}\Gamma_{en} & \Gamma_{in}-\Gamma_{en} & 0 \\ \Gamma_{en}-\Gamma_{in} & 1+\Gamma_{in}\Gamma_{en} & 0 \\ 0 & 0 & \Gamma_{in}\Gamma_{en} \end{pmatrix}. \tag{C.12}$$

Although the tensors have been expressed in a basis where $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$, the general oblique case can be found by "rotating" the above tensors with a rotation matrix, $\overline{\overline{\mathbf{S}}}$. For a given tensor $\overline{\overline{\mathbf{A}}}$ expressed for the case $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$, the new tensor $\overline{\overline{\mathbf{A'}}}$ corresponding to $\mathbf{B}_0$ making an angle θ with the positive z-axis is obtained by $\overline{\overline{\mathbf{A'}}} = \overline{\overline{\mathbf{S}}}\,\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{S}}}^{\mathsf{T}}$, where T represents the matrix transpose and the form of the rotation matrix is given by

$$\overline{\overline{\mathbf{S}}} = \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{pmatrix}. \tag{C.13}$$

The angle θ is defined by the relations $\sin(\theta) = B_{0x}/B_0$ and $\cos(\theta) = B_{0z}/B_0$, and may in general vary as a function of space: $\theta = \theta(\mathbf{r})$. Further discussion and details, including the solution for the fields in the atmosphere (not shown here) can be found in Reference [14].

Bibliography

[1] K. Akimoto, H. L. Rowland, and K. Papadopoulos. Electromagnetic radiation from strong langmuir turbulence. *Physics of Fluids*, 31(8):2185–2189, 1988.

[2] P. M. Banks and G. Kockarts. *Aeronomy: Part A*. Academic Press, New York, 1973.

[3] P. M. Banks and G. Kockarts. *Aeronomy: Part B*. Academic Press, New York, 1973.

[4] R. Barr, M. T. Rietveld, H. Kopka, P. Stubbe, and E. Nielsen. Extra-low-frequency radiation from the polar electrojet antenna. *Nature*, 317:155–157, 1985.

[5] T. F. Bell and H. D. Ngo. Electrostatic Lower Hybrid Waves Excited by Electromagnetic Whistler Mode Waves Scattering From Planar Magnetic-Field-Aligned Plasma Density Irregularities. *J. Geophys. Res.*, 95(A1):149–172, 1990.

[6] J. J. Berthelier, M. Godefroy, F. Leblanc, M. Malingre, M. Menvielle, D. Lagoutte, J. Y. Brochot, F. Colin, F. Elie, C. Legendre, P. Zamora, D. Benoist, Y. Chapuis, J. Artru, and R. Pfaff. ICE, the electric field experiment on DEMETER. *Planet. Space Sci.*, 54(5):456–471, April 2006.

[7] J.J. Berthelier, M. Godefroy, F. Leblanc, E. Seran, D. Peschard, P. Gilbert, and J. Artru. IAP, the thermal plasma analyzer on DEMETER. *Planet. Space Sci.*, 54(5):487–501, April 2006.

[8] S.I. Braginskii. Transport Properties in a Plasma. In M.A. Leontovish, editor, *Reviews of Plasma Physics*, pages 205–311. Consultants Bureau, New York, 1965.

[9] Kenneth G. Budden. *The propagation of radio waves: The theory of radio waves of low power in the ionosphere and magnetosphere*. Cambridge University Press, Cambridge, 1985.

[10] D. L. Carpenter, M. A. Spasojevic, T. F. Bell, U. S. Inan, B. W. Reinisch, I. A. Galkin, R. F. Benson, J. L. Green, S. F. Fung, and S. A. Boardsen4. Small-scale field-aligned plasmaspheric density structures inferred from the Radio Plasma Imager on IMAGE. *J. Geophys. Res.*, 107(A9):1258, 2002.

[11] G. B. Carpenter. VHF and UHF bistatic observations of a region of the ionosphere modified by a high power radio transmitter. *Radio Sci.*, 9(11):965–969, 1974.

[12] Francis F. Chen. *Introduction to Plasma Physics and Controlled Fusion, Volume 1: Plasma Physics*. Plenum Press, New York, second edition, 1984.

[13] M. B. Cohen, U. S. Inan, D. Piddyachiy, N. G. Lehtinen, and M. Gokowski. Magnetospheric injection of ELF/VLF waves with modulated or steered HF heating of the lower ionosphere. *J. Geophys. Res.*, 116(A06308), June 2011.

[14] B. Eliasson, C. L. Chang, and K. Papadopoulos. Generation of ELF and ULF electromagnetic waves by modulated heating of the ionospheric F2 region. *J. Geophys. Res.*, 117(A10320), 2012.

[15] B. Eliasson and K. Papadopoulos. Numerical study of mode conversion between lower hybrid and whistler waves on short-scale density striations. *J. Geophys. Res.*, 113(A09315), September 2008.

[16] V. L. Ginzburg. *The propagation of electromagnetic waves in plasmas*. Pergamon Press, Braunschweig, second edition, 1970.

[17] N. A. Gondarenko, S. L. Ossakow, and G. M. Milikh. Generation and evolution of density irregularities due to self-focusing in ionospheric modifications. *J. Geophys. Res.*, 110(A09304), 2005.

[18] A. V. Gurevich. Nonlinear effects in the ionosphere. *Phys. - Uspekhi*, 50(11):1091 – 1121, 2007.

[19] A. V. Gurevich, A.V. Lukyanov, and K. P. Zybin. Anomalous absorption of powerful radio waves on the striations developed during ionospheric modification. *Phys. Lett. A*, 211(6):363–372, 1996.

[20] A. V. Gurevich, G. M. Milikh, and I. S. Shluger. Nonlinear thermal focusing of radio waves in the lower ionosphere. *Geomagn. Aeron.*, 16:613–619, 1976.

[21] B. Gustavsson, M. T. Rietveld, N. V. Ivchenko, and M. J. Kosch. Rise and fall of electron temperatures: Ohmic heating of ionospheric electrons from underdense HF radio wave pumping. *J. Geophys. Res.*, 115(A12332), 2010.

[22] B. Gustavsson, T. Sergienko, M. T. Rietveld, F. Honary, A. Steen, B. U. E. Brandstrom, T. B. Leyser, A. L. Aruliah, T. Aso, M. Ejiri, and S. Marple. First tomographic estimate of volume distribution of HF-pump enhanced airglow emission. *J. Geophys. Res.*, 106(A12):29105 – 29123, 2001.

[23] F. Honary, N. Borisov, M. Beharrell, and A. Senior. Temporal development of the magnetic zenith effect. *J. Geophys. Res.*, 116(A06309), June 2011.

[24] J. D. Huba, G. Joyce, and J. A. Fedder. Sami2 is Another Model of the Ionosphere (SAMI2) - A new low-latitude ionosphere model. *J. Geophys. Res.*, 105(A10):23035–23053, 2000.

[25] Michael C. Kelley. *The Earth's Ionosphere: Plasma Physics and Electrodynamics*. Academic Press, London, second edition, 2009.

[26] H. C. Koons. Observations of large-amplitude, whistler mode wave ducts in the outer plasmasphere. *J. Geophys. Res.*, 94(A11):15393–15397, 1989.

[27] J.-P. Lebreton, S. Stverak, P. Travnicek, M. Maksimovic, D. Klinge, S. Merikallio, D. Lagoutte, B. Poirier, P.-L. Blelly, Z. Kozacek, and M. Salaquarda. The ISL Langmuir probe experiment processing onboard DEMETER: Scientific objectives, description and first results. *Planet. Space Sci.*, 54(5):472–486, April 2006.

[28] Bengt Lundborg and Bo Thidé. Standing wave pattern of HF radio waves in the ionospheric reflection region: 2. Applications. *Radio Sci.*, 21(3):486–500, 1986.

[29] R. L. Lysak. Propagation of Alfvén Waves Through the Ionosphere. *Phys. Chem. Earth*, 22(7-8):757–766, 1997.

[30] A. Mahmoudian, W. A. Scales, P. A. Bernhardt, H. Fu, S. J. Briczinski, and M. J. McCarrick. Investigation of ionospheric stimulated Brillouin scatter generated at pump frequencies near electron gyroharmonics. *Radio Sci.*, 48(6):685–697, November 2013.

[31] G. M. Milikh, A. G. Demekhov, K. Papadopoulos, A. Vartanyan, J. D. Huba, and G. Joyce. Model for artificial ionospheric duct formation due to HF heating. *Geophys. Res. Lett.*, 37(L07803), April 2010.

[32] G. M. Milikh, E. Mishin, I. Galkin, A. Vartanyan, C. Roth, and B. W. Reinisch. Ion outflows and artificial ducts in the topside ionosphere at HAARP. *Geophys. Res. Lett.*, 37(L18102), September 2010.

[33] G. M. Milikh, A. Vartanyan, K. Papadopoulos, and M. Parrot. Focusing of HF radio-waves by ionospheric ducts. *J. Atmos. Solar-Terrestrial Phys.*, 73(13):1674–1680, August 2011.

[34] Gennady Milikh, Alex Gurevich, Kiril Zybin, and Jim Secan. Perturbations of GPS signals by the ionospheric irregularities generated due to HF-heating at triple of electron gyrofrequency. *Geophys. Res. Lett.*, 35(L22102), November 2008.

[35] K. Papadopoulos. Satellite threat due to high altitude nuclear detonations. In *Present. Futur. Sp. 1st Meet. Sci. Panel, Eisenhower Inst.*, Washington, D. C., 2001.

[36] K. Papadopoulos and C. L. Chang. GENERATION OF ELF/ULF WAVES IN THE IONOSPHERE BY DYNAMO PROCESSES. *Geophys. Res. Lett.*, 12(5):279–282, 1985.

[37] K. Papadopoulos, C. L. Chang, J. Labenski, and T. Wallace. First demonstration of HF-driven ionospheric currents. *Geophys. Res. Lett.*, 38(L20107), 2011.

[38] K. Papadopoulos, C. L. Chang, P. Vitello, and A. Drobot. On the efficiency of ionospheric ELF generation. *Radio Sci.*, 25(6):1311–1320, 1990.

[39] K. Papadopoulos, N. A. Gumerov, X. Shao, I. Doxas, and C. L. Chang. HF-driven currents in the polar ionosphere. *Geophys. Res. Lett.*, 38(L12103), 2011.

[40] K. Papadopoulos, T. Wallace, M. McCarrick, G. M. Milikh, and X. Yang. On the Efficiency of ELF/VLF Generation Using HF Heating of the Auroral Electrojet. *Plasma Phys. Reports*, 29(7):561–565, July 2003.

[41] K. Papadopoulos, T. Wallace, G. M. Milikh, W. Peter, and M. McCarrick. The magnetic response of the ionosphere to pulsed HF heating. *Geophys. Res. Lett.*, 32(L13101), 2005.

[42] R. P. Perrine, G. M. Milikh, K. Papadopoulos, J. D. Huba, G. Joyce, M. Swisdak, and Y. Dimant. An interhemispheric model of artificial ionospheric ducts. *Radio Sci.*, 41(RS4002), 2006.

[43] D. Piddyachiy, U. S. Inan, T. F. Bell, N. G. Lehtinen, and M. Parrot. DEMETER observations of an intense upgoing column of ELF/VLF radiation excited by the HAARP HF heater. *J. Geophys. Res.*, 113(A10308), October 2008.

[44] M. Platino, U. S. Inan, T. F. Bell, M. Parrot, and E. J. Kennedy. DEMETER observations of ELF waves injected with the HAARP HF transmitter. *Geophys. Res. Lett.*, 33(L16101), 2006.

[45] M. Platino, U. S. Inan, T. F. Bell, J. Pickett, E. J. Kennedy, J. G. Trotignon, J. L. Rauch, and P. Canu. Cluster observations of ELF/VLF signals generated by modulated heating of the lower ionosphere with the HAARP HF transmitter. *Ann. Geophys.*, 22(7):2643–2653, July 2004.

[46] M. T. Rietveld, H. Kopka, and P. Stubbe. D-region characteristics deduced from pulsed ionospheric heating under aurora1 electrojet conditions. *J. Atmos. Terr. Phys.*, 48(4):311–326, 1986.

[47] M. T. Rietveld, M. J. Kosch, N. F. Blagoveshchenskaya, V. A. Kornienko, T. B. Leyser, and T. K. Yeoman. Ionospheric electron heating, optical emissions, and striations induced by powerful HF radio waves at high latitudes: Aspect angle dependence. *J. Geophys. Res.*, 108(A4):1141, 2003.

[48] Robert W. Schunk and Andrew F. Nagy. *Ionospheres: Physics, Plasma Physics, and Chemistry*. Cambridge University Press, Cambridge, 2000.

[49] E. Sergeev, S. Grach, A. Shindin, E. Mishin, P. Bernhardt, S. Briczinski, B. Isham, M. Broughton, J. LaBelle, and B. Watkins. Artificial Ionospheric Layers during Pump Frequency Stepping Near the 4th Gyroharmonic at HAARP. *Phys. Rev. Lett.*, 110(6):065002, February 2013.

[50] X. Shao, B. Eliasson, A. S. Sharma, G. Milikh, and K. Papadopoulos. Attenuation of whistler waves through conversion to lower hybrid waves in the low-altitude ionosphere. *J. Geophys. Res.*, 117(A04311), April 2012.

[51] X. Shao, K. Papadopoulos, and A. S. Sharma. Control of the energetic proton flux in the inner radiation belt by artificial means. *J. Geophys. Res.*, 114(A07214), 2009.

[52] M. M. Shvarts and S. M. Grach. Interaction of upper and lower hybrid waves and generation of the downshifted maximum feature of stimulated electromagnetic emissions. *J. Atmos. Solar-Terrestrial Phys.*, 59(18):2421–2429, December 1997.

[53] A. V. Streltsov, M. Lampe, W. Manheimer, G. Ganguli, and G. Joyce. Whistler propagation in inhomogeneous plasma. *J. Geophys. Res.*, 111(A03216), 2006.

[54] D. G. Swanson. *Plasma Waves*. Institute of Physics Publishing, Bristol, second edition, 2003.

[55] D. G. Swanson. *Plasma Kinetic Theory*. CRC Press, Boca Raton, 2008.

[56] A. Vartanyan, G. M. Milikh, B. Eliasson, A. C. Najmi, C. L. Chang, M. Parrot, and K. Papadopoulos. Generation of whistler waves by continuous HF heating of the upper ionosphere. *(To be submitted.)*, 2015.

[57] A. Vartanyan, G. M. Milikh, E. Mishin, M. Parrot, I. Galkin, B. Reinisch, J. Huba, G. Joyce, and K. Papadopoulos. Artificial ducts caused by HF heating of the ionosphere by HAARP. *J. Geophys. Res.*, 117(A10307), October 2012.

[58] M. Voiculescu, T. Nygrén, A. Aikio, and R. Kuula. An olden but golden EISCAT observation of a quiet-time ionospheric trough. *J. Geophys. Res.*, 115(A10315), October 2010.

[59] J. R. Woodroffe, A. V. Streltsov, A. Vartanyan, and G. M. Milikh. Whistler propagation in ionospheric density ducts: Simulations and DEMETER observations. *J. Geophys. Res. Sp. Phys.*, 118(11):7011–7018, November 2013.